

PT and diagnosis of cancers and viral diseases -
 XX
 PS claim 1; Page 80; 150pp; English.
 XX
 CC AAV4540 to AAV46214 represent specifically claimed immunogenic peptides
 CC having a human major histocompatibility complex (MHC) class I (also
 CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic
 CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes
 CC HLA A2.1, A1, A2.2 or A24.1 or HLA B or C) and induce a cytotoxic T cell
 CC response against the antigen from which the peptide is derived.
 CC cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
 CC normally induced by an antigen in the form of a peptide fragment bound
 CC to a HLA molecule, rather than the intact foreign antigen itself, and
 CC are particularly important in tumour rejection and in fighting viral
 CC infections. The peptides are therefore useful therapeutically to treat
 CC or prevent viral infections and cancers in mammals (especially humans)
 CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
 CC They can be administered as vaccines to elicit an immune response in
 CC individuals susceptible or otherwise at risk of viral infection or
 CC cancer, or used to treat chronic or acute conditions. They are also
 CC useful diagnostically, and can be used to induce a cytotoxic T cell
 CC response, by contacting a cytotoxic T cell with the peptide e.g. to
 CC produce CTLs ex vivo for infusion back into a patient. The
 CC polypeptides encoding the immunogenic peptides are also useful
 CC therapeutically and for immunisation as above.
 XX
 SO Sequence 1 AA:

 Query Match 0.0%; Score 0; DB 20; Length 1;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 YY 1 X 1
 DB 1 Y 1

 RESULT 2
 AAF10185
 ID AAF10185 standard; Protein; 2 AA.
 XX
 AC AAF10185;
 XX
 DE 14-AUG-1992 (first entry)
 XX
 DE Sequence of synthetic polypeptide for the produ. of aspartame.
 XX
 KW Aspartyl-phenylalanine methyl ester; artificial sweetener;
 XX aspartame.
 XX
 PN EP46258-A.
 XX
 PD 24 SEP-1981.
 XX
 PF 02-MAR-1981; RIIP-0300857.
 XX
 PK 14-MAR-1980; 800S-010462.
 XX
 PA (CETU) CETUS CORP.
 XX
 PI rose JF, White LD, Bahl CP;
 XX
 DR WPI: 1981-721340/40.
 DR N-PSDB; AANI0059.
 XX
 PT Aspartame produ. from synthetic polypeptide - produced by cloned
 PT microorganism
 XX
 PS Disclosure; Page 5; 18pp; English.
 XX
 CC The inventors claim a method for the produ. of aspartame by first
 CC synthesising double-stranded DNA in which a coding strand has

CC alternating codings for Asp and the Phe in sufficient number to
 CC produce a polypeptide which is stable in predetermined host
 CC microorganisms. The DNA strand is then inserted into a cloning
 CC vehicle so that the resulting chimera directs the synthesis of the
 CC Asp-Phe protein. This protein is then cleaved with CNBr or trypsin
 CC to release the peptide which, after benzylation, can be fragmented
 CC by digestion with cyanogenbromide.
 XX
 SO Sequence 2 AA:

 Query Match 0.0%; Score 0; DB 2; Length 2;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 YY 1 X 1
 DB 1 d 1

 RESULT 3
 AAF50013
 ID AAF50013 standard; Protein; 2 AA.
 XX
 AC AAF50013;
 XX
 DE 03-SEP-1991 (first entry)
 XX
 DE Sequence linking insulin B chain B(1-29) to A chain a(1-21) to form
 DE a biosynthetic insulin precursor.
 XX
 KW Insulin precursor; diabetes; proteolysis resistant.
 XX
 FE Key location/qualifiers
 FT Misc-difference 1 /note- "attached to B(1-29)"
 FT FT
 FT Misc-difference 2 /note "attached to A(1-21)"
 FT FT
 PN EP16329-A.
 PN
 PD 04-DEC-1985.
 XX
 PF 29-MAY-1985; 85EP-0116297.
 XX
 PR 08-FEB-1985; 85DK-0000582.
 PR 30-MAY-1984; 84DK-0002665.
 PR 29-MAY-1985; 85DK-0002385.
 PR 15-NOV-1990; 90EP-0121887.
 XX
 PA (NOVO) NOVO INDUSTRIAL A/S.
 XX
 PI Markussen J, Foll N, Ammerer G, Hansen M, Thim L, Norris K;
 PI Voigt BO;
 XX
 DR WPI: 1985-304970/49.
 XX
 PT DNA sequence encoding insulin precursor having correctly
 PT positioned disulphide bridges and resistant to proteolytic degradation
 XX
 PS Claim 5; Page 32; 47pp; English.
 XX
 CC The biosynthetic insulin precursors are generated largely with
 CC correctly positioned disulphide bridges between the A- and B-
 CC moieties, and are more resistant to proteolytic degradation than
 CC previously.
 XX
 SO Sequence 2 AA:

 Query Match 0.0%; Score 0; DB 6; Length 2;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

KM anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antileish-
 KM vinnery; antileishmaniasis; antibacterial; antileish; antiparasitic;
 KM cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KM neurological disease; infection; human; secreted protein.
 OS Homo sapiens.
 XX W0200058513-A1.
 PN 05-OCT-2000.
 PD 22-MAR-2000; 2000WO-0507506.
 PE 26-MAR-1999; 99US-0126507.
 PR 17-OCT-1999; 99US-0172412.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA (POSEF) POSFN C A
 XX Posen CA, Ruben SM, Komatsoulis G;
 PI WPI: 2000-544644/56.
 DR N-PSDB: AAC73913
 XX Fully time polypeptide sequences, and their amino acid sequences
 PT polypeptides, used in the treatment and diagnosis of cancers,
 PI autoimmune disorders, and skin disorders.
 XX Claim 11: Page 387; 41pp; English.
 PS The polynucleotide sequences given in AAC79740 to AAC79789 encode the
 XX human secreted proteins encoded by the genes AAC73913-73918. The genes
 CC and proteins are useful for preventing, ameliorating or treating medical
 CC conditions, e.g. by protein or gene therapy. The genes are isolated from
 CC a range of human tissues disclosed in the specification. The nucleic
 CC acids, proteins, antibodies and (ant)ibodies are used in the
 CC diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
 CC ovarian cancer, and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC hemolytic anemia, autoimmune thyroiditis, diabetes mellitus,
 CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and osteoporosis;
 CC (c) cardiovascular disorders such as myocardial ischemias, and
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Sequence 2 AA:
 SQ
 Query Match 0.0%; Score 0; DB 21; Length 2;
 Best local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 X 1
 Db 1 p 1
 RESULT 9
 AAB44709
 ID AAB44709 standard; Protein; 2 AA.
 AC AAB44709;
 XX 12-FEB-2001 (first entry)
 DE Human secreted protein sequence encoded by gene 15 SEQ ID NO:75.
 XX Human secreted protein; diagnosis; anti-allergic; immunosuppressive;
 KM anti-rheumatic; antiproliferative; cytostatic; cardiac; vasotropic;
 KM cerebroprotective; neuroprotective; neuroprotective; antibacterial; virucide;
 KM fungicide; ophthalmological; vinnery; gene therapy; autoimmune disease;

KM hyperproliferative disorder; cardiovascular disorder; angiogenesis;
 KM cerebrovascular disorder; nervous system disorder; infection;
 KM wound healing; food additive; preservative; skin aging.
 OS Homo sapiens.
 XX W0200058494-A1.
 PN 05-OCT-2000.
 PD 23-MAR-2000; 2000WO-0507578.
 PE 26-MAR-1999; 99US-0126507.
 PR 07 JAN 2000; 2000US-0174672.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Posen CA, Ruben SM, Komatsoulis G;
 PI WPI: 2000-544644/56.
 DR N-PSDB: AAC79754.
 XX New nucleic acid molecules encoding 50 human secreted proteins for
 PT diagnosis, prevention, treating or ameliorating medical conditions and
 PI used as food additives or preservatives.
 XX Claim 11: Page 346; 37pp; English.
 PS The polynucleotide sequences given in AAC79740 to AAC79789 encode the
 XX human secreted proteins given in AAB44695 to AAB44744. AAB44745 to
 CC AAB44760 represent human secreted polypeptide sequences and proteins
 CC homologous to them, which are given in the exemplification of the present
 CC invention. Human secreted proteins have activities based on the tissues
 CC and cells the genes are expressed in. Examples of activities include:
 CC antiarthritic; immunosuppressive; antirheumatic; antiproliferative;
 CC cytostatic; cardiac; vasotropic; neuroprotective; neurotrophic;
 CC cerebroprotective; antibacterial; virucide; fungicide; ophthalmological;
 CC and vinnery. The polypeptides and polypeptides can be used to
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
 CC in diagnosing a pathological condition or susceptibility to a
 CC pathological condition. Disorders which are diagnosed or treated include
 CC autoimmune diseases, hyperproliferative disorders, cardiovascular
 CC disorders, cerebrovascular disorders, angiogenesis, nervous system
 CC disorders, infections caused by bacteria, viruses and fungi and organ
 CC disorders. The polypeptides can also be used to aid wound healing and
 CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
 CC maintain organs before transplantation, for supporting cell culture of
 CC primary tissues, to regenerate tissues and in chemotaxis. The
 CC polypeptide can also be used as a food additive or preservative to
 CC increase or decrease storage capabilities, fat content, lipid, protein,
 CC carbohydrate, vitamins, minerals, cofactors and other nutritional
 CC components. AAC79741 to AAC79789 and AAB44745 to AAB44760 represent sequences used in
 CC the exemplification of the present invention.
 CC Sequence 2 AA:
 SQ
 Query Match 0.0%; Score 0; DB 21; Length 2;
 Best local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 X 1
 Db 1 m 1
 RESULT 10
 AAB44925
 ID AAB44925 standard; Protein; 2 AA.
 AC AAB44925;
 XX

DT 12-FEB-2001 (first entry)
 XX
 DE Human secreted protein encoded by gene 9.
 XX
 KW Secreted protein; human; immunosuppressive; antiarthritic; antirheumatic;
 KW antiproliferative; cytosolic; cardiant; vasotropic; cerebroprotective;
 KW neutropic; neuroprotective; antibacterial; virucide; fungicide;
 KW ophthalmological; gene therapy; treatment; autoimmune disease; infection;
 KW cardiovascular disorder; nervous system disorder; ocular disorder;
 KW wound healing; epithelial cell proliferation; skin aging; mental state;
 KW transplantation; metabolism modulation.
 XX
 OS Homo sapiens.
 XX
 PN W0200055200-A1.
 XX
 DE 21-SEP-2000.
 XX
 PE 09-MAR-2000; 2000MC-0506042.
 XX
 PR 12-MAR-1999; 99US-0124143.
 PR 03-DEC-1999; 99US-0168663.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM, Komatsoulis G;
 PI
 DE WPI: 2000-656008/63.
 DE N-PSDB: AAC79906.
 XX
 PT Isolated human secretory proteins, nucleic acids encoding them and
 PT antibodies directed against them, useful for diagnosing and treating
 PT disorders related to the proteins such as cancer, Alzheimer's disease
 PT and Parkinsons -
 XX
 PS Claim 11: Page 483; 453pp; English.
 XX
 CC This invention describes a novel isolated polypeptide (I) and its
 CC encoding nucleic acid molecule (II) which have immunosuppressive,
 CC antiarthritic, antirheumatic, antiproliferative, cytosolic, cardiant,
 CC vasotropic, cerebroprotective, footropic, neuroprotective, antibacterial,
 CC virucide, fungicide and ophthalmological activity and which can be used
 CC for gene therapy. (I) and (II) are used to prevent, treat or ameliorate
 CC a medical condition in e.g. humans, mice, rabbits, goats, horses, cats,
 CC dogs, chickens or sheep. (I) and (II) are also used in diagnosing a
 CC pathological condition or susceptibility to a pathological condition. The
 CC antibodies to (I) can also be used in alleviating symptoms associated
 CC with the disorders and in diagnostic immunoassays e.g. radioimmunoassays
 CC or enzyme linked immunosorbent assays (ELISA) disorders which are
 CC diagnosed or treated include autoimmune diseases e.g. rheumatoid
 CC arthritis, hyperproliferative disorders e.g. neoplasms of the breast or
 CC liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular
 CC disorders e.g. cerebral ischemia, angiodenosis, nervous system disorders
 CC e.g. Alzheimer's disease, infections caused by bacteria, viruses and
 CC fungi and ocular disorders e.g. corneal infection. The polypeptides can
 CC also be used to aid wound healing and epithelial cell proliferation, to
 CC prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, for supporting cell culture of primary tissues, to
 CC regenerate tissues and in chemotherapy. The polypeptides are used to
 CC modulate mammalian metabolism, to change human's mental state or
 CC physical state by influencing biohybrid circadian rhythms, depression
 CC tendency for violence, tolerance for pain, reproductive capabilities,
 CC hormonal or endocrine levels, appetite, libido, memory, stress or other
 CC cognitive qualities, as a food additive or preservative, such as to
 CC increase or decrease storage capabilities, fat content, lipid, protein,
 CC carbohydrates, vitamins, minerals, cofactors or other nutritional
 CC components.
 XX
 SO Sequence 2 AA;
 Query Match 0 0% Score 0 PR 21 Length 2;
 Best Local Similarity 0 0% Pseq No 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 X 1
 DB 1 1 1
 RESULT 11
 ID AAB44951
 ID AAB44951 standard, Protein, 2 AA.
 XX
 AC AAB44951;
 XX
 DT 12-FEB-2001 (first entry)
 DE
 DE Human secreted protein encoded by gene 35.
 XX
 KW Secreted protein; human; immunosuppressive; antiarthritic; antirheumatic;
 KW antiproliferative; cytosolic; cardiant; vasotropic; cerebroprotective;
 KW neutropic; neuroprotective; antibacterial; virucide; fungicide;
 KW ophthalmological; gene therapy; treatment; autoimmune disease; infection;
 KW cardiovascular disorder; nervous system disorder; ocular disorder;
 KW wound healing; epithelial cell proliferation; skin aging; mental state;
 KW transplantation; metabolism modulation.
 XX
 OS Homo sapiens.
 XX
 PN W0200055200-A1.
 XX
 DE 21-SEP-2000.
 XX
 PE 09-MAR-2000; 2000MC-0506042.
 XX
 PR 12-MAR-1999; 99US-0124143.
 PR 03-DEC-1999; 99US-0168663.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM, Komatsoulis G;
 PI
 DE WPI: 2000-656008/63.
 DE N-PSDB: AAC79932.
 XX
 PT Isolated human secretory proteins, nucleic acids encoding them and
 PT antibodies directed against them, useful for diagnosing and treating
 PT disorders related to the proteins such as cancer, Alzheimer's disease
 PT and Parkinsons -
 XX
 PS Claim 11: Page 333; 453pp; English.
 XX
 CC This invention describes a novel isolated polypeptide (I) and its
 CC encoding nucleic acid molecule (II) which have immunosuppressive,
 CC antiarthritic, antirheumatic, antiproliferative, cytosolic, cardiant,
 CC vasotropic, cerebroprotective, footropic, neuroprotective, antibacterial,
 CC virucide, fungicide and ophthalmological activity and which can be used
 CC for gene therapy. (I) and (II) are used to prevent, treat or ameliorate
 CC a medical condition in e.g. humans, mice, rabbits, goats, horses, cats,
 CC dogs, chickens or sheep. (I) and (II) are also used in diagnosing a
 CC pathological condition or susceptibility to a pathological condition. The
 CC antibodies to (I) can also be used in alleviating symptoms associated
 CC with the disorders and in diagnostic immunoassays e.g. radioimmunoassays
 CC or enzyme linked immunosorbent assays (ELISA) disorders which are
 CC diagnosed or treated include autoimmune diseases e.g. rheumatoid
 CC arthritis, hyperproliferative disorders e.g. neoplasms of the breast or
 CC liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular
 CC disorders e.g. cerebral ischemia, angiodenosis, nervous system disorders
 CC e.g. Alzheimer's disease, infections caused by bacteria, viruses and
 CC fungi and ocular disorders e.g. corneal infection. The polypeptides can
 CC also be used to aid wound healing and epithelial cell proliferation, to
 CC prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, for supporting cell culture of primary tissues, to
 CC regenerate tissues and in chemotherapy. The polypeptides are used to
 CC modulate mammalian metabolism, to change human's mental state or

CC physical state by interacting lipoproteins, circadian rhythm, depression
CC tendency for violence following the pain, reproductive capabilities,
CC hormonal or endocrine levels, appetite, libido, memory, stress or other
CC cognitive qualities, as a food additive or preservative, such as to
CC increase or decrease storage capabilities, fat content, lipid, protein,
CC carbohydrates, vitamins, minerals, cofactors or other nutritional
CC components.

XX Sequence 2 AA.

Query Match 0.08; Score 0; DB 21; Length 2;
Best local Similarity 0.08; Pct. Neg. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1
DB 1 1 1

RESULT 12

AA27796 AAB27796 standard; Protein, 2 AA.

AC AAB27796;

DT 29-JAN-2001 (first entry)

XX Human secreted protein #3.

DE Human secreted protein #3.

XX Cytostatic; immunosuppressive; neotropic; neuroprotective; antiviral;

KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;

KW valeric; anticonvulsant; antibacterial; antifungal; antiparasitic;

KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;

KW neurological disease; infection; human; secreted protein.

XX Homo sapiens.

OS Homo sapiens.

FR W020005199 A1.

XX 21-SEP-2000

PD 09-MAR-2000; 2000W0-0806014

XX 12-MAR-1999; 99US-0124095

XX 11-JUN-1999; 99US-0138598

XX 03-DEC-1999; 99US-0168665

XX (HUMA-) HUMAN GENOME SCI INC.

PA Rosen CA, Ruben SM, Komatsoulis G;

XX WPI: 2000-572359/53.

XX N-PSDB: AAC59217.

DR Isolated nucleic acid molecule encoding a human secreted protein is

XX used in preventing, treating or ameliorating a medical condition

PT Claim 11; Page 371-372, 433pp, English.

PS Sequences AAB27794-827840 represent the amino acid sequences of 47

XX human secreted proteins encoded by the genes AAC59215-059261. The genes

CC and proteins are useful for preventing, ameliorating or treating medical

CC conditions, e.g. by protein or gene therapy. The genes are isolated from

CC a range of human tissues disclosed in the specification. The nucleic

CC acids, proteins, antibodies and (ant)agonists are useful in the

CC diagnosis, treatment and prevention of: (a) cancer, e.g. breast and

CC ovarian cancer, and other cancers of the adrenal gland, bone, bone

CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;

CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune

CC hemolytic anemia, autoimmune thyroiditis, diabetes mellitus,

CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative

CC colitis; (c) cardiovascular disorders such as myocardial ischemias;

CC colitis; (d) infectious diseases such as viral, bacterial, fungal

CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and

CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal

CC and parasitic infections.

XX Sequence 2 AA;

Query Match 0.08; Score 0; DB 21; Length 2;
Best local Similarity 0.08; Pct. Neg. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1
DB 1 1 1

RESULT 13

AA27827 AAB27827 standard; Protein, 2 AA.

AC AAB27827;

DT 29-JAN-2001 (first entry)

XX Human secreted protein #34.

DE Human secreted protein #34.

XX Cytostatic; immunosuppressive; neotropic; neuroprotective; antiviral;

KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;

KW valeric; anticonvulsant; antibacterial; antifungal; antiparasitic;

KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;

KW neurological disease; infection; human; secreted protein.

XX Homo sapiens.

OS Homo sapiens.

FR W020005199 A1.

XX 21-SEP-2000

PD 09-MAR-2000; 2000W0-0806014

XX 12-MAR-1999; 99US-0124095

XX 11-JUN-1999; 99US-0138598

XX 03-DEC-1999; 99US-0168665

XX (HUMA-) HUMAN GENOME SCI INC.

PA Rosen CA, Ruben SM, Komatsoulis G;

XX WPI: 2000-572359/53.

XX N-PSDB: AAC59248.

DR Isolated nucleic acid molecule encoding a human secreted protein is

XX used in preventing, treating or ameliorating a medical condition

PT Claim 11; Page 385; 433pp; English.

PS Sequences AAB27794-827840 represent the amino acid sequences of 47

XX human secreted proteins encoded by the genes AAC59215-059261. The genes

CC and proteins are useful for preventing, ameliorating or treating medical

CC conditions, e.g. by protein or gene therapy. The genes are isolated from

CC a range of human tissues disclosed in the specification. The nucleic

CC acids, proteins, antibodies and (ant)agonists are useful in the

CC diagnosis, treatment and prevention of: (a) cancer, e.g. breast and

CC ovarian cancer, and other cancers of the adrenal gland, bone, bone

CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;

CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune

CC hemolytic anemia, autoimmune thyroiditis, diabetes mellitus,

CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative

CC colitis; (c) cardiovascular disorders such as myocardial ischemias;

CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and

CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal

CC and parasitic infections.

SU Sequence 2 AA:

Query Match 0.0%; Score 0; DB 21; Length 2;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 DB 1 1 1

RESULT 14
 AAB27830
 ID AAB27830 standard; Protein: 2 AA.

AC AAB27830;
 DE 29-JAN-2001 (first entry)
 DE Human secreted protein #37.
 DE Cytostatic; immunosuppressive; neurotropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KW valeric; anticonvulsant; antibacterial; antitumor; antiparasitic;
 KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.

OS Homo sapiens.
 PN W020005199-A1
 IN 21-SEP-2000

DE 09-MAR-2000; 2000W0-0506014
 PR 12-MAR-1999; 9905-0124095.
 PR 11-JUN-1999; 9905-0138598.
 PR 01-DEC-1999; 9905-0168665.

PA (HUMAN-) HUMAN GENOME SCI INC.
 PI Rosen CA, Ruben SM, Komatsoulis G;

DR WPI: 2000-572359/53.
 DR N-PSDB; AAC59251.

PT Isolated nucleic acid molecule encoding a human secreted protein is
 used in preventing, treating or ameliorating a medical condition.

Claim 11: Page 386; 433pp; English.

Sequences AAB27794-B27840 represent the amino acid sequences of 47
 human secreted proteins encoded by the genes AAC52215-C52261. The genes
 and proteins are useful for preventing, ameliorating or treating medical
 conditions, e.g. by protein or gene therapy. The genes are isolated from
 a range of human tissues disclosed in the specification. The nucleic
 acids, proteins, antibodies and (ant)agonists are useful in the
 diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
 ovarian cancer, and other cancers of the adrenal gland, bone, bone
 marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,
 Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 and parasitic infections.

SU Sequence 2 AA:

Query Match 0.0%; Score 0; DB 21; Length 2;

Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 DB 1 m 1

RESULT 15
 AAB27834
 ID AAB27834 standard; Protein: 2 AA.

AC AAB27834;
 DE 29-JAN-2001 (first entry)
 DE Human secreted protein #41.
 DE Cytostatic; immunosuppressive; neurotropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KW valeric; anticonvulsant; antibacterial; antitumor; antiparasitic;
 KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.

OS Homo sapiens.
 PN W020005199-A1.
 IN 21-SEP-2000.

DE 09-MAR-2000; 2000W0-0506014
 PR 12-MAR-1999; 9905-0124095.
 PR 11-JUN-1999; 9905-0138598.
 PR 01-DEC-1999; 9905-0168665.

PA (HUMAN-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM, Komatsoulis G;

DR WPI: 2000-572359/53.
 DR N-PSDB; AAC59255.

PT Isolated nucleic acid molecule encoding a human secreted protein is
 used in preventing, treating or ameliorating a medical condition.

Claim 11: Page 387; 433pp; English.

Sequences AAB27794-B27840 represent the amino acid sequences of 47
 human secreted proteins encoded by the genes AAC52215-C52261. The genes
 and proteins are useful for preventing, ameliorating or treating medical
 conditions, e.g. by protein or gene therapy. The genes are isolated from
 a range of human tissues disclosed in the specification. The nucleic
 acids, proteins, antibodies and (ant)agonists are useful in the
 diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
 ovarian cancer, and other cancers of the adrenal gland, bone, bone
 marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,
 Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 and parasitic infections.

SU Sequence 2 AA:

Query Match 0.0%; Score 0; DB 21; Length 2;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 m 1

Search completed: September 30, 2001, 10:31:24
Job time: 154 sec

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

Genotools version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd

OM protein - protein search, using sw model

Run on: September 30, 2001, 10:30:07 : Search time 21.36 Seconds

(without alignments)
14,265 Million cell updates/sec

Title: US-09-498-556-85

Perfect score: 4

Sequence: 1 XXXX 4

Scoring table: BLAST/MSM62

Gapop 10.0, Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

1: p1r1:
2: p1r2:
3: p1r3:
4: p1r4:

pred. No. is the number of results predicted by BLAST. The higher the score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	DB ID	Description
1	0	0.0	3	PHITDO
2	0	0.0	3	thyroliberin - Bom
3	0	0.0	3	thyroliberin - pig
4	0	0.0	3	thyroliberin - she
5	0	0.0	3	thyroliberin - eas
6	0	0.0	3	growth-modulating
7	0	0.0	3	bursin - chicken
8	0	0.0	3	spinal cord peptid
9	0	0.0	3	spinal cord peptid
10	0	0.0	3	thyroliberin releas
11	0	0.0	3	thyroliberin releas
12	0	0.0	3	thyroliberin releas
13	0	0.0	3	thyroliberin releas
14	0	0.0	3	thyroliberin releas
15	0	0.0	3	thyroliberin releas
16	0	0.0	3	thyroliberin releas
17	0	0.0	3	thyroliberin releas
18	0	0.0	3	thyroliberin releas
19	0	0.0	3	thyroliberin releas
20	0	0.0	3	thyroliberin releas
21	0	0.0	3	thyroliberin releas
22	0	0.0	3	thyroliberin releas
23	0	0.0	3	thyroliberin releas
24	0	0.0	3	thyroliberin releas
25	0	0.0	3	thyroliberin releas
26	0	0.0	3	thyroliberin releas
27	0	0.0	3	thyroliberin releas
28	0	0.0	3	thyroliberin releas
29	0	0.0	3	thyroliberin releas

30	0	0.0	4	PL0146	carbon-monoxide de
31	0	0.0	4	A37832	phenol 2-methoxyge
32	0	0.0	4	A48360	gamma subunit of p
33	0	0.0	4	A40697	biotin A - Citoba
34	0	0.0	4	A61300	22K superhelical D
35	0	0.0	4	A57745	D-mannanase hydrol
36	0	0.0	4	A41890	Protein D - Escher
37	0	0.0	4	A43014	hypothetical prote
38	0	0.0	4	D41654	hypothetical prote
39	0	0.0	4	B43848	cell surface adhes
40	0	0.0	4	A40505	hypothetical prote
41	0	0.0	4	A40870	phospholipase C (E
42	0	0.0	4	A40804	endothucanase P -
43	0	0.0	4	T46627	hypothetical prote
44	0	0.0	4	S09478	globulin IV alpha
45	0	0.0	4	S53508	starvation induced

ALIGNMENTS

RESULT 1
thyroliberin - Bombina orientalis
C:Species: Bombina orientalis
C:Date: 31-Dec-1991 #seqrev-revision 12-May-1995 #text-change 20-Mar-1998
C:Accession: A0919; A01415
C:Author: T. Nakajima, T. Rayasuhara, T. Nakajima, T. Chem. Pharm. Bull. 23, 3301-3303, 1975
A:Title: Occurrence of thyroliberin-Pro-NH-2 in the frog skin.
A:Reference number: A0919; MIMD:76138399
A:Accession: A0919
A:Molecule type: protein
A:Residues: 1-3 (YAS)
C:Superfamily: thyroliberin precursor
C:Keywords: amidated carboxyl end, glutamoyl gluc, hormone; pyroglutamic acid
F1/Modified site: pyroglutamic carboxylic acid (Glu) #status experimental
F3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match Similarity 0.0%; Score 0; DB 2; Length 3;
Best local 0; prod No 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Q7 1 X 1
DB 1 Q 1

RESULT 2
thyroliberin - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 23-Oct-1981 #seqrev-revision 12-May-1995 #text-change 20-Mar-1998
C:Accession: A01415
C:Author: P. M. G. Barrett, J. E. Roberts, C. Y. Schally, A. V. Biochem. Biophys. Res. Commun. 97, 706-710, 1981
A:Title: The identity of chemical and hormonal properties of the thyroliberin P-1-Asin
A:Reference number: A01415; MIMD:7003904
A:Accession: A01415
A:Molecule type: protein
A:Residues: 1-3 (KAS)
F:Author: J. E. Roberts, C. Y. Schally, A. V. Biochem. Biophys. Res. Commun. 97, 706-710, 1981
A:Title: The identity of chemical and hormonal properties of the thyroliberin P-1-Asin
A:Reference number: A01415; MIMD:7003904
A:Accession: A01415
A:Contents: annotation
A:Name: thyroliberin and P-1-Asin and P-1-Asin
C:Superfamily: thyroliberin precursor
C:Keywords: amidated carboxyl end, hormone; pyroglutamic acid
F1/Modified site: pyroglutamic carboxylic acid (Glu) #status experimental
F3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0% Score 0; DB 2; Length 3;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

UY 1 X 1
DB 1 G 1

RESULT 4

RESHI

Thyroliberin - sheep

C:Species: ovis orientalis aries, ovis aries (domestic sheep)

C>Date: 31-Dec-1991 #sequence_revision: 12 Mar, 1995 #text_change: 20 Mar, 1996

C:Accession: A94750; A01415

C:ReleaseDate: J.F., D.M.; Burgess, R.; Dunn, T.F.; Vale, W.; Guillemin, R.; Ward, D.N.

C:Mass Spectrum: 5; 221-228, 1971

A:Title: The elucidation of the primary structure of the hypothalamic thyroid stimulating

A:Reference number: A94750

A:Accession: A94750

A:Molecule type: protein

A:Residues: 1-4 <HS>

A:Note: R.; Dunn, T.F.; Desiderio, D.; Ward, D.N.; Vale, W.; Guillemin, R.

A:Title: Characterization of ovine hypothalamic hypophysiotropic TSH-releasing factor.

A:Reference number: A94161; M01D:7016386

A:Contents: annotation

A:Note: physicochemical characteristics and biological activities of the natural and syn-

C:Superfamily: thyroliberin precursor

C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid

F3/Modified site: pyroglutamic carboxyl end (Glu) #status experimental

F3/Modified site: amidated carboxyl end (Pro) #status experimental

UY 1 X 1

DB 1 G 1

Query Match 0.0% Score 0; DB 2; Length 3;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

UY 1 X 1
DB 1 G 1

RESULT 4
A92971
Thyroliberin - eastern newt (tentative sequence)

C:Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)

C>Date: 31-Dec-1991 #sequence_revision: 31-Dec-1991 #text_change: 20-Mar-1996

C:Accession: A92971; A01415

C:ReleaseDate: J.F., D.M.; Burgess, R.; Dunn, T.F.

C:Mass Spectrum: 25; 471-478, 1974

A:Title: Biosynthesis of thyroliberin releasing factor by newt (*Triturus viridescens*) brd

A:Reference number: A92971; M01D:75035605

A:Accession: A92971

A:Molecule type: protein

A:Residues: 1-3 <HS>

A:Note: a peptide with the chromatographic and electrophoretic characteristics of thyrol-

iberin or glutamic acid

C:Superfamily: thyroliberin precursor

C:Keywords: amidated carboxyl end; calyculase; fluid; hormone; hypothalamus; pyroglutamic

F3/Modified site: pyroglutamic carboxyl end (Glu) #status experimental

F3/Modified site: amidated carboxyl end (Pro) #status experimental

UY 1 X 1

DB 1 G 1

Query Match 0.0% Score 0; DB 2; Length 3;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 5
GRHO
growth modulating peptide - human

C:Species: Homo sapiens (man)

C>Date: 22-Oct-1981 #sequence_revision: 24 May 1995 #text_change: 20 Jun 2000

C:Accession: A01421

C:ReleaseDate: D.H.; Pickart, L.; Thaler, M.M.

C:Mass Spectrum: 35; 324-325, 1977

A:Title: Growth-modulating serum tripeptide is glyxyl-histidyl-lysine.

A:Reference number: A01421; M01D:77162469

A:Accession: A01421

A:Molecule type: protein

A:Residues: 1-3 <SCM>

A:Note: this serum tripeptide is found to stimulate growth of some cell types and to

C:Superfamily: unassigned animal peptides

Query Match 0.0% Score 0; DB 2; Length 3;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

UY 1 X 1
DB 1 G 1

RESULT 6

A60898
bursin - chicken

C:Species: Gallus gallus (chicken)

C>Date: 31-Dec-1993 #sequence_revision: 31-Dec-1993 #text_change: 18 Aug 2000

C:Accession: A60898

C:ReleaseDate: R.; Audhya, T.; Kroon, D.; Heavner, G.; Viamonte, G.; Goldstein, G.

C:Mass Spectrum: 231; 997-999, 1986

A:Title: Tripeptide structure of bursin, a selective B-cell-differentiating hormone o

A:Reference number: A60898; M01D:8612916

A:Accession: A60898

A:Molecule type: protein

A:Residues: 1-3 <AUD>

A:Superfamily: unassigned animal peptides

C:Keywords: amidated carboxyl end; hormone

F3/Modified site: amidated carboxyl end (Gly) #status experimental

UY 1 X 1

DB 1 K 1

Query Match 0.0% Score 0; DB 2; Length 3;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 7

A23751
spinal cord peptide SCP-4 - pig

C:Species: Sus scrofa domestica (domestic pig)

C>Date: 28-Sep-1987 #sequence_revision: 28-Sep-1987 #text_change: 18-Aug-2000

C:Accession: A23751

C:ReleaseDate: R.; Hsi, K.; Chou, R.-L.; Chou, C.-T.; Chang, H.-L.; Lu, Y.-A.; Guo, S.-Y.; Wu, S.-X.; Tsou

Arch. Biochem. Biophys. 240, 178-184, 1985

A:Reference number: A23751; M01D:85250425

A:Accession: A23751

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-3 <HS1>

C:Superfamily: unassigned animal peptides

Query Match 0.0% Score 0; DB 2; Length 3;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1
DB 1 Q 1

RESULT 8

R23751
spinal cord peptide SCP-5 - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 28-Sep-1987 #sequence_revision 28 Sep 1987 #text_change 18 Aug 2000
C:Accession: R23751
R:Hist., K.L.: Chen, R.L.: Chen, Z.G.: Zhang, H.L.: Lu, Y.A.: Guo, S.Y.: Wu, S.X.: Tsou, K.
Arch. Biochem. Biophys. 240, 178-183, 1985
A:Reference number: A23751, MIM:95250125
A:Accession: R23751
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-4-His
C:Superfamily: unassigned animal peptides

Query Match
Best local similarity 0.0% Score 0; DB 2; Length 3;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1
DB 1 M 1

RESULT 9

A13802
thyrotropin-releasing hormone-like peptide - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 27-Feb-1990 #sequence_revision 31-Dec-1992 #text_change 18-Aug-2000
C:Accession: A13802
R:Cockle, S.M.: Aitken, A.: Bay, F.: Smyth, D.G.
J. Biol. Chem. 264, 7798-7799, 1989
A:Title: A novel peptide, pyroglutamylglutaminylproline amide, in the rabbit prostate comp
A:Reference number: A13802; MIM:89255136
A:Accession: A13802
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-3-COO
C:Superfamily: unassigned animal peptides
C:Keywords: inhibitor; antioxy; and; pyroglutamic acid
E1/Modified site: pyroglutamic acid (Gln) #status experimental
E3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match
Best local similarity 0.0% Score 0; DB 2; Length 3;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1
DB 1 Q 1

RESULT 10

A22565
R-phycoerythrin alpha-1 chain - red alga (Gastrocionium coulteri) (fragment)
C:Species: Gastrocionium coulteri
C:Date: 07-Mar-1988 #sequence_revision 07-Mar-1988 #text_change 23-Mar-1993
C:Accession: A22565
R:Klotz, A.V.: Glazer, A.N.
J. Biol. Chem. 260, 4856-4863, 1985
A:Title: Characterization of the bilin attachment sites on R-phycoerythrin
A:Reference number: A22565; MIM:85183601
A:Accession: A22565
A:Molecule type: protein
A:Residues: 1-4-Phe

Query Match
Best local similarity 0.0% Score 0; DB 2; Length 3;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1
DB 1 Q 1

RESULT 11

P00010
angiotensin converting enzyme inhibitor (FIP-3) common fig
R:Alvarado Lamas, I.: Luis latex peptide 3
C:Species: Ficus carica (common fig)
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 08-Dec-1995
C:Accession: P00010
R:Maruyama, S.: Miyoshi, S.: Tanaka, H.
Agric. Biol. Chem. 52, 2763-2767, 1989
A:Title: Angiotensin I-converting enzyme inhibitors derived from Ficus carica.
A:Reference number: P00010
A:Accession: P00010
A:Molecule type: protein
A:Residues: 1-3-CHAr
A:Experimental source: latex
C:Keywords: angiotensin-converting enzyme inhibitor

Query Match
Best local similarity 0.0% Score 0; DB 2; Length 3;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1
DB 1 L 1

RESULT 12

S13894
histidyl dehydrogenase (EC 1.1.1.22) wild cabbage (fragment)
C:Species: Brassica oleracea (wild cabbage)
C:Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 06-Jun-1997
C:Accession: S13894
R:Nadal, A.: Scheidegger, A.
Arch. Biochem. Biophys. 280, 127-132, 1991
A:Title: Purification and characterization of histidyl dehydrogenase from cabbage.
A:Reference number: S13894; MIM:9112784
A:Accession: S13894
A:Molecule type: protein
A:Residues: 1-3-CHAr
A:Experimental source: var capitata
C:Keywords: dimer; NAD; oxidoreductase

Query Match
Best local similarity 100.0% Score 100; DB 2; Length 3;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 X 1
DB 1 X 1

RESULT 13

A43391
TRH-like tripeptide - alfalfa
C:Species: Medicago sativa (alfalfa)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: A43391
R:Lackey, D.B.
J. Biol. Chem. 267, 17508-17511, 1992
A:Title: Isolation and structural determination of a novel TRH-like tripeptide from

A:Reference number: A43391; MUID:92388092

A:Accession: A4391

A:Molecule type: protein

A:Residues: 1-3 sLMC

C:Keywords: amidated carboxyl end; pyrrolutamic acid

F1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F1/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match

Best Local Similarity 0.0%; Score 0; DB 2; Length 3;

Best Local Similarity 0.0%; Pred. No: 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

DB 1 Q 1

RESULT 14

E47196

bradykinin-potentiating peptide 5 - island jararaca

C:Species: Bothrops jarararus (island jararaca)

C>Date: 14-Feb-1992 #sequence_revision 01-Dec-1992 #text_change 05-Aug-1994

C:Accession: E47196

R:Pintra, A.C.O.; Vieira, C.A.; Gagliolo, J.R.

J. Protein Chem. 9, 221-227, 1990

A:Title: Primary structure and biological activity of bradykinin potentiating peptides

A:Reference number: A47196; MUID:90351557

A:Accession: E47196

A>Status: preliminary

A:Molecule type: protein

C:Residues: 1-3 sLMC

C:Keywords: pyrrolutamic acid

F1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match

Best Local Similarity 0.0%; Score 0; DB 2; Length 3;

Best Local Similarity 0.0%; Pred. No: 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

DB 1 Q 1

RESULT 15

E47196

bradykinin-potentiating peptide 6 - island jararaca

C:Species: Bothrops jarararus (island jararaca)

C>Date: 14-Feb-1992 #sequence_revision 01-Dec-1992 #text_change 05-Aug-1994

C:Accession: E47196

R:Pintra, A.C.O.; Vieira, C.A.; Gagliolo, J.R.

J. Protein Chem. 9, 221-227, 1990

A:Title: Primary structure and biological activity of bradykinin potentiating peptides

A:Reference number: A47196; MUID:90351557

A:Accession: E47196

A>Status: preliminary

A:Molecule type: protein

C:Residues: 1-3 sLMC

C:Keywords: pyrrolutamic acid

F1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match

Best Local Similarity 0.0%; Score 0; DB 2; Length 3;

Best Local Similarity 0.0%; Pred. No: 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

DB 1 Q 1



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd

QM protein - protein search, using SW model

Run on: September 30, 2001, 10:29:34 Search time: 13.87 seconds
(without alignments)
9.879 Million cell updates/sec

Title: US-09-498-556-85

Perfect score: 4

Sequence: 1 XXXX 4

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_39*

pred No. is the number of results predicted by change to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length DB	ID	Description
1	0	0.0	3	1 GRWM_HUMAN	P01157 homo sapien
2	0	0.0	3	1 TMF_VIBRI	P24272 vibrio fischeri
3	0	0.0	3	1 THYL_PIG	P01151 sus scrofa
4	0	0.0	4	1 ACHL_ACHER	P35404 acharia fa
5	0	0.0	4	1 DCM1_PSECH	P19416 pseudomonas
6	0	0.0	4	1 DCM5_PSECH	P19918 pseudomonas
7	0	0.0	4	1 EOSTI_HUMAN	P02731 homo sapien
8	0	0.0	4	1 PAP3_HIPME	P42562 hitudo medi
9	0	0.0	4	1 PAP4_HIPME	P42561 hitudo medi
10	0	0.0	4	1 FLPE_HIPME	P01162 macrocallis
11	0	0.0	4	1 FMRE_MAVNI	P36515 saccharomy
12	0	0.0	4	1 PMO1_YEAST	P01858 homo sapien
13	0	0.0	4	1 TPTI_HUMAN	P81817 carinus ma
14	0	0.0	5	1 AL14_CARMA	P13071 citrobacter
15	0	0.0	5	1 R10A_CITTEP	P12677 salmonella
16	0	0.0	5	1 R10A_SALTY	P12697 citrobacter
17	0	0.0	5	1 R10R_CITTEP	P12678 salmonella
18	0	0.0	5	1 R10R_SALTY	P30425 bolitropis in
19	0	0.0	5	1 BPP7_BOTIN	P41853 atroposth
20	0	0.0	5	1 FARP_AKTR	P81864 pardachius
21	0	0.0	5	1 FARP_PARMA	P01373 periplaneta
22	0	0.0	5	1 PRGT_PERAM	P19991 acheta dome
23	0	0.0	5	1 SUGA_ACHDO	P54714 canis famill
24	0	0.0	5	1 TPIS_CANEA	P13973 escherichia
25	0	0.0	5	1 TRM3_BCOLI	P35288 escherichia
26	0	0.0	5	1 TRF2_MATIF	P34634 mus musculu
27	0	0.0	5	1 UPN1_MOUSE	P34635 mus musculu
28	0	0.0	5	1 UX44_CHITS	P25554 mytilus edu
29	0	0.0	6	1 ACPL_PARIT	P13737 mytilus edu
30	0	0.0	6	1 CIP1_MYTED	P13737 mytilus edu
31	0	0.0	6	1 CIP2_MYTED	P13737 mytilus edu
32	0	0.0	6	1 FARP_MONFX	P41866 molecule ex
33	0	0.0	6	1 LOK1_LOOMI	P41491 locusta mig

34	0	0.0	6	1 OVM_LEPDE	P42985 leptinotars
35	0	0.0	6	1 TMF_SARBU	P41495 sarcophaga
36	0	0.0	6	1 TRF1_PSHPI	P4414 sarcophaga
37	0	0.0	6	1 UN06_CLOPA	P81351 clostridium
38	0	0.0	6	1 VP19_HSVIK	P23210 herpes simp
39	0	0.0	7	1 ALL2_CARMA	P81805 carinus ma
40	0	0.0	7	1 ALL3_CARMA	P81806 carinus ma
41	0	0.0	7	1 ALL4_CARMA	P81807 carinus ma
42	0	0.0	7	1 ALL5_CARMA	P81808 carinus ma
43	0	0.0	7	1 ALL7_CYDPO	P82158 cydia pomon
44	0	0.0	7	1 CARP_MYED	P10420 mytilus edu
45	0	0.0	7	1 UN06_PINPS	P81675 pinus pinas

ALIGNMENTS

RESULT 1
GRWM_HUMAN STANDARD: PRT: 3 AA.
ID GRWM_HUMAN
AC P01157
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last annotation update)
DE GROWTH-MODULATING PEPTIDE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eularchia; Primates; Catarrhini; Hominoidea; Homo.
RN NCBI_Taxid:9606;
RN 111
RP SEQUENCE.
RX MEDLINE:77162369; Pubmed 858356;
PA Schlesinger P H, Probst L, Thaler M.M.;
BT "Growth-modulating signal tripeptide is Glycyl-histidyl-lysine";
EL Experientia 33:324-325(1977).
CC -I- MISCELLANEOUS: THIS SERUM TRIPEPTIDE HAS BEEN FOUND TO STIMULATE
CC GROWTH OF SOME CELL TYPES AND TO INHIBIT OTHER TYPES IN VITRO.
DR PIR: A01421; GKHU.
SQ SEQUENCE 3 AA: 340 MW: 6331810000000000 GPC64;

Query Match 0.0% Score 0.0 PRT: Length 3;
Host Local Similarity 0.0% Pred No 0;
Matches 0; Conservatives 0; Mismatches 1; Indels 0; Gaps 0;
UY 1 X 1
DB 1 G 1
RESULT 2
LUXE_VIBRI STANDARD: PRT: 3 AA.
ID LUXE_VIBRI
AC P24272;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE LONG-CHAIN-FATTY-ACID-LIOCLIPERIN-COMPONENT LIGASE (EC 6.2.1.19)
DE (ACYL-PROTEIN SYNTHETASE) (FRAGMENT).
GN LUXE.
OS Vibrio fischeri.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_Taxid:668;
RN 111
RP SEQUENCE FROM N.A.
RX MEDLINE:31972326; Pubmed 2251256;
PA Swartzman F, Kapoor S, Graham A.F., Melnyk E.A.;
BT "A new Vibrio fischeri lux gene encodes a bidirectional termination
BT site for the lux operon";
EL J Bacteriol 172:6797-6802(1990).
CC FUNCTION: AYL PROTEIN SYNTHETASE ACTIVATES TETRAHYDROIC ACID.
CC IT IS A COMPONENT OF THE FATTY ACID REDUCTASE COMPLEX RESPONSIBLE
CC FOR CONVERTING TETRADECANOIC ACID TO THE ALDEHYDE WHICH SERVES AS

CC SUBSTRATE IN THE PHOSPHATASE-CATALYZED REACTION.
 CC -1- CATALYTIC ACTIVITY: ATP + AN ACID + PROTEIN - AMP + PYROPHOSPHATE
 CC + AN ACTI-PROTEIN THIOESTER.
 CC -1- PATHWAY: SECOND STEP IN THE BIOLUMINESCENT FATTY ACID REDUCTASE
 CC COMPLEX.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/annoucy>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M28122; NOT_ANNOTATED_CDS.
 CC Bioluminescence: Lipase.
 CC E1: M28122; 1
 CC SEQUENCE: 4 AA; 374 MW; 6AA3404000000000 CPO64;
 CC -----
 CC Query Match: 0.0%; Score 0; DB 1; Length 3;
 CC Best Local Similarity: 0.0%; Prod. No. 0;
 CC Matches: 0; Conservative: 0; Mismatches: 1; Indels: 0; Gaps: 0;
 CC -----
 CC QY 1 X 1
 CC Db 1 1 1
 CC -----
 CC RESULT 4
 CC THYL_PIG STANDARD; PRT; 4 AA.
 CC AC P0151;
 CC 21-JUN-1986 (Rel. 01, Created)
 CC 21-JUN-1986 (Rel. 01, Last sequence update)
 CC 15-DEC-1998 (Rel. 37, Last annotation update)
 CC DE THYROLIBERIN (THYROTROPIN-RELEASING HORMONE) (TRH).
 CC SUS scrota (Pig), ovis aries (Sheep),
 CC US Bombina orientalis (Oriental fire-bellied toad), and
 CC OS Melophthalmus viridescens (Eastern newt) (Triturus viridescens).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
 CC OX NCBI_TaxID:9823, 9940, 8146, 816;
 CC RN [1]
 CC SEQUENCE:
 CC SPECIES: Pig; TISSUE: Hypothalamus;
 CC RX MEDLINE:70146150; PubMed:4984938;
 CC RA Nair R.M.G., Barrett J.F., Howers C.Y., Schally A.V.;
 CC 15-JUN-1994 (Rel. 29, Last sequence update)
 CC 15-JUN-1994 (Rel. 36, Last annotation update)
 CC DE ACHALIN-T.
 CC US Achalina fulica (Giant African snail).
 CC OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
 CC OC Achatinacea; Achatinidae; Achatina.
 CC OX NCBI_TaxID:6530;
 CC RN [1]
 CC SEQUENCE: CHARACTERIZATION, AND SYNTHESIS.
 CC STRAIN:FERUSSAC; TISSUE:Ganlion;
 CC RX MEDLINE:89273551; PubMed:2597281;
 CC RA Kametani Y., Minakata H., Keny P.T.M., Iwashita T., Watanabe K.,
 CC Nawa K., Sun X.P., Yousufi A., Kim K.H., Novales-Li P.,
 CC PA Novales E.T., Kanapi C.G., Takeuchi H., Kemoto K.;
 CC 17-APR-1994 (Rel. 29, Last sequence update)
 CC 17-APR-1994 (Rel. 36, Last annotation update)
 CC DE Achatina fulica, an endogenous neurosecretory tetrapeptide from Achatina
 CC fulica ferrussac containing a D-amino acid residue.
 CC RL Biochem. Biophys. Res. Commun. 160:1015-1020(1989).
 CC RN [2]
 CC CHARACTERIZATION.
 CC STRAIN:FERUSSAC; TISSUE:Heart atrium;
 CC RX MEDLINE:91264856; PubMed:1675568;
 CC RA Fujimoto K., Kubota I., Yasuda-Kanetani Y., Minakata H., Muroto K.,
 CC KA Yoshida M., Hatada A., Kaneoka Y., Kobayashi M.;
 CC 17-APR-1994 (Rel. 29, Last sequence update)
 CC 17-APR-1994 (Rel. 36, Last annotation update)
 CC DE Achatina fulica, and its possible function.
 CC RL Biochem. Biophys. Res. Commun. 177:847-853(1991).
 CC RN [3]

FT "Characterization of ovine hypothalamic hypophysiotropic
 RT TRH-releasing factor."
 RL Nature 226:321-325(1970).
 RL [5]
 RP SEQUENCE.
 RC SPECIES: B.orientalis; TISSUE: SKIN;
 RX MEDLINE:76138399; PubMed:815011;
 RA Yasuhara T., Nakajima T.;
 RT "Isolation and characterization of thyrotropin-releasing factor."
 RL J. Neurochem. 23:471-478(1974).
 CC -1- FUNCTION: TRH FUNCTIONS AS A REGULATOR OF THE BIOSYNTHESIS OF TRH
 CC IN THE ANTERIOR PITUITARY GLAND AND AS A NEUROTRANSMITTER/
 CC NEUROMODULATOR IN THE CENTRAL AND PERIPHERAL NERVOUS SYSTEMS.
 CC DE PIR: A01415; RHPGT.
 CC DR PIR: A93750; RHPGT.
 CC DR PIR: A93750; RHPGT.
 CC DR PIR: A92971; A92971.
 CC KW Amidation.
 CC FT MOD_RES 1 1 PYROGLUTAMINE CARBOXYLIC ACID.
 CC FT MOD_RES 3 4 AMIDATION.
 CC SQ SEQUENCE 3 AA; 380 MW; 7761F6H000000000 CPO64;
 CC -----
 CC Query Match: 0.0%; Score 0; DB 1; Length 3;
 CC Best Local Similarity: 0.0%; Prod. No. 0;
 CC Matches: 0; Conservative: 0; Mismatches: 1; Indels: 0; Gaps: 0;
 CC -----
 CC QY 1 X 1
 CC Db 1 0 1
 CC -----
 CC RESULT 4
 CC ACHI_ACHFU STANDARD; PRT; 4 AA.
 CC AC P35904;
 CC 01-JUN-1994 (Rel. 29, Created)
 CC 01-JUN-1994 (Rel. 29, Last sequence update)
 CC 15-JUN-1994 (Rel. 36, Last annotation update)
 CC DE ACHALIN-T.
 CC US Achalina fulica (Giant African snail).
 CC OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
 CC OC Achatinacea; Achatinidae; Achatina.
 CC OX NCBI_TaxID:6530;
 CC RN [1]
 CC SEQUENCE: CHARACTERIZATION, AND SYNTHESIS.
 CC STRAIN:FERUSSAC; TISSUE:Ganlion;
 CC RX MEDLINE:89273551; PubMed:2597281;
 CC RA Kametani Y., Minakata H., Keny P.T.M., Iwashita T., Watanabe K.,
 CC Nawa K., Sun X.P., Yousufi A., Kim K.H., Novales-Li P.,
 CC PA Novales E.T., Kanapi C.G., Takeuchi H., Kemoto K.;
 CC 17-APR-1994 (Rel. 29, Last sequence update)
 CC 17-APR-1994 (Rel. 36, Last annotation update)
 CC DE Achatina fulica, an endogenous neurosecretory tetrapeptide from Achatina
 CC fulica ferrussac containing a D-amino acid residue.
 CC RL Biochem. Biophys. Res. Commun. 160:1015-1020(1989).
 CC RN [2]
 CC CHARACTERIZATION.
 CC STRAIN:FERUSSAC; TISSUE:Heart atrium;
 CC RX MEDLINE:91264856; PubMed:1675568;
 CC RA Fujimoto K., Kubota I., Yasuda-Kanetani Y., Minakata H., Muroto K.,
 CC KA Yoshida M., Hatada A., Kaneoka Y., Kobayashi M.;
 CC 17-APR-1994 (Rel. 29, Last sequence update)
 CC 17-APR-1994 (Rel. 36, Last annotation update)
 CC DE Achatina fulica, and its possible function.
 CC RL Biochem. Biophys. Res. Commun. 177:847-853(1991).
 CC RN [3]

RP X-RAY CRYSTALLOGRAPHY.
 RX MEDLINE=93014529; PubMed 139265;
 RA Ishida T., In Y., Doi M., Yasuda-Kamatani Y., Minakata H.,
 RA Iwashita T., Nemoto K.;
 RT "Crystal structure and molecular conformation of acetalin-I
 RT (H₂O)-D-Phe-Ala-Asp-CH₃ on endogenous neurotransmitter containing a
 RT D-amino acid residue."
 RL Int. J. Pept. Protein Res. 50:258-264(1992)
 CC -1- FUNCTION: NEUROEXCITATORY PEPTIDE; INCREASES THE IMPULSE FREQUENCY
 CC AND PROGRESSES A SPIKE REWARDING OF THE IDENTIFIED HEART EXCITATORY
 CC NEURON (PUN); ALSO ENHANCES THE AMPLITUDE AND FREQUENCY OF THE
 CC HEART BEAT. HAS ALSO AN EFFECT ON SYMPATHIC OTHER MUSCLES.
 DR PIR: A32480; A32480;
 KW Hormone; D-amino acid;
 FT MOD_RES 2
 FT NON_TER 2
 SQ SEQUENCE 4 AA; 408 MW; 6AADD9C810000000 CWC64;
 Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 X 1
 DB 1 G 1

RESULT 5
 ID DCMS_PSECH STANDARD; PRT; 4 AA.
 AC P19916;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, last sequence update)
 DT 01-DEC-1992 (Rel. 24, last annotation update)
 DE CARBON MONOXIDE DEHYDROGENASE LARGE CHAIN (EC 1.2.99.2) (FRAGMENT).
 OS Pseudomonas carboxydohydrogenasa
 CC Bacteria; Proteobacteria;
 CX NCBI_Taxid=290;
 RN (1)
 RP SEQUENCE:
 RX MEDLINE=90055678; PubMed 2818128;
 RA Kraut M., Hugendieck L., Herwig S., Meyer O.;
 RT "Hemology and distribution of co dehydrogenase structural genes in
 RT carboxydohydrophic bacteria."
 RL Arch. Microbiol. 152:335-341(1989).
 CC -1- CATALYTIC ACTIVITY: CO + H(2)O + ACCEPTOR -> CO(2) + REDUCED
 CC ACCEPTOR.
 CC -1- COFACTOR: MOLYBDENUM.
 CC -1- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
 CC SMALL.
 DR PIR: P10140; P10140
 KW Oxidoreductase; Molybdenum;
 FT NON_TER 4
 FT SEQUENCE 4 AA; 441 MW; 7761P8760000000 CWC64;
 Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 X 1
 DB 1 M 1

RESULT 6
 ID DCMS_PSECH STANDARD; PRT; 4 AA.
 AC P19918;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, last sequence update)
 DT 01-DEC-1992 (Rel. 24, last annotation update)
 DE CARBON MONOXIDE DEHYDROGENASE SMALL CHAIN (EC 1.2.99.2) (FRAGMENT)
 OS Pseudomonas carboxydohydrogenasa

OS Pseudomonas carboxydohydrogenasa.
 CC Bacteria; Proteobacteria;
 CX NCBI_Taxid=290;
 RN (1)
 RP SEQUENCE:
 RX MEDLINE=90055678; PubMed 2818128;
 RA Kraut M., Hugendieck L., Herwig S., Meyer O.;
 RT "Hemology and distribution of co dehydrogenase structural genes in
 RT carboxydohydrophic bacteria."
 RL Arch. Microbiol. 152:335-341(1989).
 CC -1- CATALYTIC ACTIVITY: CO + H(2)O + ACCEPTOR -> CO(2) + REDUCED
 CC ACCEPTOR.
 CC -1- COFACTOR: MOLYBDENUM.
 CC -1- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
 CC SMALL.
 DR PIR: P10146; P10146;
 KW Oxidoreductase; Molybdenum;
 FT NON_TER 4
 FT SEQUENCE 4 AA; 420 MW; 6B933900P0000000 CWC64;
 Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 X 1
 DB 1 M 1

RESULT 7
 ID EOST_HUMAN STANDARD; PRT; 4 AA.
 AC P02731;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, last sequence update)
 DT 21-JUL-1986 (Rel. 01, last annotation update)
 DE EOSTINOPHILINOTACTIC PEPTIDES.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo
 CX NCBI_Taxid=9606;
 RN (1)
 RP SEQUENCE:
 RX MEDLINE=7607412; PubMed 1060093;
 RA Goetzel E.G., Austen K.F.;
 RT "Purification and synthesis of eosinophilic tetrapeptides of
 RT human lung tissue: identification as eosinophil chemotactic factor of
 RT anaphylaxis."
 RL Proc. Natl. Acad. Sci. U.S.A. 72:4123-4127(1975).
 CC -1- MISCELLANEOUS: THESE PEPTIDES ARE RELEASED FROM MAST CELLS IN LUNG
 CC (AND OTHER TISSUES) DURING HYPERSENSITIVITY REACTIONS
 CC (ANAPHYLAXIS). THEIR ACTIVITIES, PREFERENTIALLY AFFECTING
 CC EOSINOPHILS, INCLUDE CHEMOTAXIS, CHEMOTACTIC DEACTIVATION, RELEASE
 CC OF ENZYMES, AND STIMULATION OF THE HEXOSE MONOPHOSPHATE SHUNT.
 DR PIR: A03190; ETHOL.
 FT VARIANT 1 V -> A (IN OTHER PEPTIDE).
 FT SEQUENCE 4 AA; 300 MW; 6B059B2A00000000 CWC64;
 Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 X 1
 DB 1 V 1

RESULT 8
 ID FAP4_HIRME STANDARD; PRT; 4 AA.
 ID FAP4_HIRME

Query Match	0.08;	Score 0;	DR 1;	Length 4;
Best Local	Similarity 0.08;	Pred. No. 0;		
Matches	0;	Conservative	0;	Mismatches 1;
				Indexis 0;
Gapis				0;
OY	I X I			
Ob	I Y I			
RESULT 10				
FILE_HIRME				
ID	FILE_HIRME	STANDARD;	PRT;	4 AA
AC	P42561;			
DI	01-NOV-1995 (Rel. 32, Created)			
DI	01-NOV-1995 (Rel. 32, Last sequence update)			
DI	01-NOV-1995 (Rel. 32, Last annotation update)			
DE	FILEFAMIDE;			

FP SEQUENCE, AND CHARACTERIZATION
 RC SPECIES=M.ombosa; TISSUE=Gallium;
 vx MEDLINE=78012038; PubMed=909875;
 RA Price D.A., Greendberg M.J.,
 RT "Purification and characterization of a cardiorespiratory neuropeptide
 RT from the central ganglia of a bivalve mollusc."
 RI Pref. Biochem. 7:261-281(1977)
 RN [3]
 RP SEQUENCE:
 RC SPECIES=N.virens;
 EX MEDLINE=90259856; PubMed=242392;
 RA Krajinak K.G., Price D.A.;
 RT "Antibiotic FMRFamide is present in the polychaete Nereis virens."
 RL Peptides 11:75-77(1990).
 RN [4]
 RP SEQUENCE:
 RC SPECIES=H.medicinalis;
 RP

EX MEDLINE-92195954; PubMed 1686943;
 RA Evans B.D., Pohl J., Karimatis M.A., Calabrese K.L.;
 RT "Identification of Human Neuropeptides in the medicinal leech."
 RL Peptides 12:897-908(1991).
 RN [5]
 RP SEQUENCE.
 KC SPECIES-H. LITREOLUS; TISSUE-Kidney;
 RA MEDLINE-9425417; PubMed-7919138;
 RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;
 RT "Peptide-related peptides from the kidney of the snail, Helisoma
 trivolvis."
 RL Peptides 15:31-36(1994).
 CC -1- FUNCTION: MONOACTIVE, GABAB-RECEPTARY SUBSTANCE PHARMACOLOGICAL
 CC ACTIVITIES: INTRACELLULAR AMPLIFICATION, INHIBITION, AND REGULATION OF
 CC CARDIAC CONTRACTION.
 CC -1- SIMILARITY: BELONGS TO THE FAPP (PMPAMIDE RELATED PEPTIDE)
 CC FAMILY.
 DR PIR: A01426; ECKR;
 DR PIR: A60418; A60418.
 DR Neuropeptide; Amidation.
 FT MOD_RES 4 4 AMIDATION.
 SO SEQUENCE 4 AA: 600 MW: 694069400000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 DB 1 F 1

RESULT 12
 RM01_YEAST
 ID RM01_YEAST STANDARD; PRT: 4 AA.
 AC P36515;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE MITOCHONRIAL 60S PROTEOMAL PROTEIN L1 (YMC1) (FRAGMENT).
 GN MRPL1.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.
 CX NCBI_TaxId=4932;
 RN [1]
 RP SEQUENCE.
 RA MEDLINE-91265156; PubMed 20627;
 RA Grahman L., Gradaak H.-R., Kraft V., Choll T., Goldschmidt-Reisin S.,
 RA Kitakawa M.;
 RT "Extended N-terminal sequencing of proteins of the large ribosomal
 subunit from yeast mitochondria."
 RL FEBS Lett. 284:51-56(1991).
 DR PIR: S17255; S17255.
 DR SGD: L0002641; MRPL1.
 KW Ribosomal protein; Mitochondrion.
 FT NON_TER 4 4
 SO SEQUENCE 4 AA: 402 MW: 777103500000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 DB 1 S 1

RESULT 13
 TUF1_HUMAN
 ID TUF1_HUMAN STANDARD; PRT: 4 AA.

AC P01858;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-JUL-2000 (Rel. 40, Last annotation update)
 DE PHAGOCYTOSIS-STIMULATING PEPTIDE (TUF1SN).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Canidia; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE.
 RA MEDLINE=72187087; PubMed=4112769;
 RA Nishida K., Constantopoulos A., Salch P.S., Najjar V.A.;
 RT "The characteristics, isolation and synthesis of the phagocytosis
 stimulating peptide tufisin."
 RL Biochem. Biophys. Res. Commun. 47:172-179(1972).
 RN [2]
 RP IMMUNOGLOBULIN CLASS.
 RA MEDLINE=68091045; PubMed=4169272;
 RA Fidaig B.V., Najjar V.A.;
 RT "The physiological role of the lymphoid system VI The stimulatory
 effect of leucophilin gamma globulin (leucokinin) on the phagocytic
 activity of human polymorphonuclear leucocyte."
 RL Biochemistry 6:3386-3392(1967).
 CC -1- MISCELLANEOUS: AN IGG (CALLED LEUCOKININ) BINDS REVERSIBLY TO THE
 CC CELL MEMBRANE OF NEUTROPHILS IN THE BLOOD. LEUCOKININASE ON THE
 CC MEMBRANE RELEASES THE ACTIVE PEPTIDE TUF1SN FROM THE GAMMA CHAIN.
 CC TUF1SN IS ESSENTIAL FOR MAXIMUM STIMULATION OF THE PHAGOCYTIC
 CC ACTIVITY OF NEUTROPHILS.
 DR PIR: A02147; A02147.
 DR MIM: 191150;
 SO SEQUENCE 4 AA: 501 MW: 7417632100000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 DB 1 T 1

RESULT 14
 AL14_CARMA
 ID AL14_CARMA STANDARD; PRT: 5 AA.
 AC P81817;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE CARCINOSTATIN 14.
 OS Carcinus maenas (Common shore crab) (Crustacea)
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
 OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
 CX NCBI_TaxId=6759;
 RN [1]
 RP SEQUENCE.
 FC TISSUE=Cerebral ganglion, and Thoracic ganglion.
 RA MEDLINE=98121193; PubMed=9461295;
 RA Duve H., Johnson A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
 RA Thorpe A.;
 RT "Isolation and identification of multiple neuropeptides of the
 RT allatostatin superfamily in the shore crab Carcinus maenas."
 RL Eur. J. Biochem. 250:727-734(1997).
 CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 KW Neuropeptide; Amidation; Multigene family.
 FT MOD_RES 5 5 AMIDATION (POTENTIAL).
 SO SEQUENCE 5 AA: 586 MW: 6126746500000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 5;

Best Local Similarity 0.08; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1

DB 1 Y 1

RESULT 15

ID BIOA_CITR STANDARD PRT 5 AA.

AC P13071;

BT 01-JAN-1990 (Rel. 13, Created)

DT 01-JAN-1990 (Rel. 13, last sequence update)

UF 01-OCT-2000 (Prt. 40, last annotation update)

DE ADENOSYLMETHIONINE-8-AMINO-7-OXONONANOATE AMINOTRANSFERASE

DE (EC 2.6.1.62) (7.8-DIAMINO-PICARONIC ACID AMINOTRANSFERASE) (DAPA

DE AMINOTRANSFERASE) (FRAGMENT).

CN BIOA.

OS Citrobacter freundii.

OC Bacteria; Proteobacteria; gamma subdivision, Enterobacteriaceae,

OC Citrobacter.

OX NCBI_TaxID:546;

RN 111

RP SEQUENCE FROM N.A.

KX MEDLINE 89066280; PubMed-2971595;

KA Shiuan D., Campbell A.;

KT "Transcriptional regulation and gene arrangement of Escherichia coli,

KT Citrobacter freundii and Salmonella typhimurium biotin operons.";

KL Gene 67:203-211(1988).

CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + 8-AMINO-7-

CC OXONONANOATE = S-ADENOSYL-4-METHYLTHIO-2-OXOBUTANOATE +

CC 7.8-DIAMINONONANOATE.

CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.

CC -1- PATHWAY: BIOTIN BIOSYNTHESIS.

CC -1- SUBUNIT: HOMODIMER.

CC -1- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT

CC AMINOTRANSFERASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: M21922; -; NOT_ANNOTATED_CDS.

DR INTERPRO: IPR000954; -;

DR PROSITE: PS00660; AA_TRANSFPR_CLASS_3; PARTIAL.

KW Biotin biosynthesis; Transferase; Aminotransferase;

KW Pyridoxal phosphate.

FT NON_TER 5

ST SEQUENCE 5 AA; 582 MW; 6AABAB1A6F00000 CRC64;

Query Match 0.08; Score 0, DB 1, length 5,

Best Local Similarity 0.08; Pred. No. 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1

DB 1 M 1

Search completed: September 30, 2001, 10:29:16
Job time: 26 sec

Geneset version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

om protein - protein search, using sw model

Run on: September 30, 2001, 10:29:37 ; Search time 18.45 seconds
(without alignments)
4,488 Million cell updates/sec

Title: US-09-498-556-85

Perfect score: 4

Sequence: 1 XXXX 4

Scoring table:

Matrix: BLOSUM62
Gapop: 10.0 ; Gapext: 0.5

Searched: 157339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 157339

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match: 0%
Maximum Match: 100%
Listing first 45 summaries

Database: Issuel-Patents-AA*

1: US-07-920-154A-12
2: US-07-791-2120-24
3: US-07-791-2120-10
4: US-08-174-055A-57
5: US-07-789-913-24
6: US-07-789-913-25
7: US-08-049-794-24
8: US-08-049-794-25
9: US-08-423-027-12
10: US-08-448-606-4
11: US-07-869-943-16
12: US-08-293-150A-24
13: US-08-293-150A-40
14: US-08-496-847-23
15: US-08-496-847-25
16: US-08-742-774-24
17: US-08-742-774-25
18: US-08-575-354-24
19: US-08-575-354-25
20: US-08-667-554A-12
21: US-08-965-918-24
22: US-08-965-918-25
23: US-09-138-439-24
24: US-09-138-439-25
25: US-08-480-640A-12
26: US-08-613-407A-24
27: US-08-613-407A-25

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	0	0.0	1	US-07-920-154A-12	Sequence 12, Appl
2	0	0.0	1	US-07-791-2120-24	Sequence 24, Appl
3	0	0.0	1	US-07-791-2120-10	Sequence 10, Appl
4	0	0.0	1	US-08-174-055A-57	Sequence 57, Appl
5	0	0.0	1	US-07-789-913-24	Sequence 24, Appl
6	0	0.0	1	US-07-789-913-25	Sequence 25, Appl
7	0	0.0	1	US-08-049-794-24	Sequence 24, Appl
8	0	0.0	1	US-08-049-794-25	Sequence 25, Appl
9	0	0.0	1	US-08-423-027-12	Sequence 12, Appl
10	0	0.0	1	US-08-448-606-4	Sequence 4, Appl
11	0	0.0	1	US-07-869-943-16	Sequence 16, Appl
12	0	0.0	1	US-08-293-150A-24	Sequence 24, Appl
13	0	0.0	1	US-08-293-150A-40	Sequence 40, Appl
14	0	0.0	1	US-08-496-847-23	Sequence 23, Appl
15	0	0.0	1	US-08-496-847-25	Sequence 25, Appl
16	0	0.0	1	US-08-742-774-24	Sequence 24, Appl
17	0	0.0	1	US-08-742-774-25	Sequence 25, Appl
18	0	0.0	1	US-08-575-354-24	Sequence 24, Appl
19	0	0.0	1	US-08-575-354-25	Sequence 25, Appl
20	0	0.0	1	US-08-667-554A-12	Sequence 12, Appl
21	0	0.0	1	US-08-965-918-24	Sequence 24, Appl
22	0	0.0	1	US-08-965-918-25	Sequence 25, Appl
23	0	0.0	1	US-09-138-439-24	Sequence 24, Appl
24	0	0.0	1	US-09-138-439-25	Sequence 25, Appl
25	0	0.0	1	US-08-480-640A-12	Sequence 12, Appl
26	0	0.0	1	US-08-613-407A-24	Sequence 24, Appl
27	0	0.0	1	US-08-613-407A-25	Sequence 25, Appl

28	0	0.0	1	US-08-801-092-10	Sequence 10, Appl
29	0	0.0	1	US-08-801-092-17	Sequence 17, Appl
30	0	0.0	1	US-08-801-092-24	Sequence 24, Appl
31	0	0.0	1	US-08-801-092-31	Sequence 31, Appl
32	0	0.0	1	US-08-801-092-38	Sequence 38, Appl
33	0	0.0	1	US-08-801-092-45	Sequence 45, Appl
34	0	0.0	1	US-09-298-017-23	Sequence 23, Appl
35	0	0.0	1	US-09-298-017-25	Sequence 25, Appl
36	0	0.0	1	US-08-295-802-12	Sequence 12, Appl
37	0	0.0	1	US-09-392-979A-22	Sequence 22, Appl
38	0	0.0	1	US-09-392-979A-25	Sequence 25, Appl
39	0	0.0	1	US-09-103-663-16	Sequence 16, Appl
40	0	0.0	1	US-08-488-237A-12	Sequence 12, Appl
41	0	0.0	1	PCT-US91-04492-4	Sequence 4, Appl
42	0	0.0	1	PCT-US91-00124-12	Sequence 12, Appl
43	0	0.0	2	US-07-679-052A-6	Sequence 6, Appl
44	0	0.0	2	US-07-914-280-1	Sequence 1, Appl
45	0	0.0	2	PCT-US93-06625-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-07-820-154A-12
Sequence 12, Application US/97820154A
Patent No. 5,626,255
GENERAL INFORMATION:
APPLICANT: Cochran Ph.D., Mark D
ATTORNEY: Junker M.S., David E
TITLE: INVENTION OF A POLYMERIC VIRUS
PCT: F. 40
SERIAL: 40
ADDRESS: 6000 S. 10th St., White
City, New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM pc compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: Patrolin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
ATTORNEY REFERENCE: 007,920,154A
FILING DATE: 19920113
CLASSIFICATION: 124
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
TELEPHONE: (212)977-0550
TELEFAX: (212)664-0525
TELEX: 422523
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-820-154A-12

Query Match 0.0% Score 0; DB 1; Length 1;

Best Local Similarity 0.0% Prod. No. 0; Model 0;

Database: us-09-498-556-85; Alignment: 1; Gaps: 0;

RESULT 2

US-07-791-2140-24
 ? Sequence 24, Application US/077912130
 ? Patent No. 5409895
 ? GENERAL INFORMATION:
 ? APPLICANT: MORISITA, Hiideaki
 ? APPLICANT: KANAMORI, Toshiohri
 ? APPLICANT: MOTOHARA, Masahito
 ? TITLE OF INVENTION: POLYPEPTIDE, DNA FRAGMENT ENCODING THE
 ? TITLE OF INVENTION: SAME AND PROCESS FOR PROTECTING THE SAME, AND ENZYME
 ? TITLE OF INVENTION: INHIBITION PROCESS, LONG COMPOSITION AND METHODS OF
 ? TITLE OF INVENTION: TREATING USING THE SAME
 ? NUMBER OF SEQUENCES: 108
 ? CORRESPONDENCE ADDRESS:
 ? ADDRESS: P.O. Box 1404
 ? CITY: Alexandria
 ? STATE: Virginia
 ? COUNTRY: United States
 ? ZIP: 22313-1404
 ? COMPUTER READABLE FORM:
 ? MEDIUM TYPE: Floppy disk
 ? COMPUTER: IBM PC compatible
 ? OPERATING SYSTEM: PC-DOS/MS-DOS
 ? SOFTWARE: Patent in release #1.0, Version #1.25
 ? CURRENT APPLICATION DATA:
 ? APPLICATION NUMBER: US/07791, 2130
 ? FILING DATE: 13-NOV-1991
 ? CLASSIFICATION: 435
 ? PRIOR APPLICATION DATA:
 ? APPLICATION NUMBER: JP 2-306745
 ? FILING DATE: 13-NOV-1990
 ? ATTORNEY/AGENT INFORMATION:
 ? NAME: Meulh, Donna M
 ? REGISTRATION NUMBER: 36,607
 ? REFERENCE/DOCKET NUMBER: 029650-032
 ? TELECOMMUNICATION INFORMATION:
 ? TELEPHONE: (703) 836-6620
 ? TELEFAX: (703) 836-6620
 ? INFORMATION FOR SEQ ID NO: 24:
 ? SEQUENCE CHARACTERISTICS:
 ? LENGTH: 1 amino acids
 ? TYPE: amino acid
 ? TOPOLOGY: linear
 ? MULTIPLE TYPE: peptide
 ? US-07-791-2140-24

Query Match 0.0%; Score 0; DB 1; Length 1;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 Db 1 C 1

RESULT 4
 US-07-791-2140-40
 ? Sequence 40, Application US/077912140
 ? Patent No. 5409895
 ? GENERAL INFORMATION:
 ? APPLICANT: MORISITA, Hiideaki
 ? APPLICANT: KANAMORI, Toshiohri
 ? APPLICANT: MOTOHARA, Masahito
 ? TITLE OF INVENTION: POLYPEPTIDE, DNA FRAGMENT ENCODING THE
 ? TITLE OF INVENTION: SAME AND PROCESS FOR PROTECTING THE SAME, AND ENZYME
 ? TITLE OF INVENTION: INHIBITION PROCESS, LONG COMPOSITION AND METHODS OF
 ? TITLE OF INVENTION: TREATING USING THE SAME
 ? NUMBER OF SEQUENCES: 108
 ? CORRESPONDENCE ADDRESS:
 ? ADDRESS: P.O. Box 1404
 ? CITY: Alexandria

STATE: Virginia
 COUNTRY: United States
 ZIP: 22313-1404
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07791, 2140
 FILING DATE: 13-NOV-1991
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 2-306745
 FILING DATE: 13-NOV-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Meulh, Donna M
 REGISTRATION NUMBER: 36,607
 REFERENCE/DOCKET NUMBER: 029650-032
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 836-6620
 TELEFAX: (703) 836-2021
 INFORMATION FOR SEQ ID NO: 40:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-07-791-2140-40

Query Match 0.0%; Score 0; DB 1; Length 1;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 Db 1 C 1

RESULT 4
 US-08-174-365A-57
 ? Sequence 57, Application US/08174365A
 ? Patent No. 5478809
 ? GENERAL INFORMATION:
 ? APPLICANT: Seiochi TANIDA et al.
 ? TITLE OF INVENTION: 4-THIAHEPTANOIC ACID DERIVATIVES
 ? NUMBER OF SEQUENCES: 106
 ? CORRESPONDENCE ADDRESS:
 ? ADDRESS: Wendtoll, Lind & Iwata
 ? STREET: 805 Fifteenth Street, N.W., #700
 ? CITY: Washington
 ? STATE: D.C.
 ? COUNTRY: U.S.A.
 ? ZIP: 20005
 ? COMPUTER READABLE FORM:
 ? MEDIUM TYPE: Diskette, 3.5 inch, 144 mb
 ? COMPUTER: IBM compatible
 ? OPERATING SYSTEM: MS-DOS
 ? SOFTWARE: Wordperfect 5.1
 ? CURRENT APPLICATION DATA:
 ? APPLICATION NUMBER: US/08/174, 365A
 ? FILING DATE: December 28, 1993
 ? CLASSIFICATION: 514
 ? PRIOR APPLICATION DATA:
 ? APPLICATION NUMBER:
 ? FILING DATE:
 ? ATTORNEY/AGENT INFORMATION:
 ? NAME: Warren M. Cheek, Jr.
 ? REGISTRATION NUMBER: 33,367
 ? REFERENCE/DOCKET NUMBER:
 ? TELECOMMUNICATION INFORMATION:
 ? TELEPHONE: 202-371-8850

TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 1 amino acid
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: modified site
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: Note: "Xaa is modified amino acid as
OTHER INFORMATION: described in specification"
US-08-174-365A-97

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 X 1
DB 1 X 1

RESULT 5
US-07-789-913-23
Sequence 23, Application US/07789913
Patent No. 555095
GENERAL INFORMATION:
APPLICANT: Miljanich, George P.
APPLICANT: Homeros, Stephen S.
APPLICANT: Fox, James A.
APPLICANT: Valentino, Karen L.
APPLICANT: Bittner, Robert S.
APPLICANT: Yamashiro, Donald H.
TITLE OF INVENTION: Delayed Treatment Method of Reducing
TITLE OF INVENTION: Ischemia Related Neuronal Damage
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Peter Dehlinger
STREET: 450 Cambridge Avenue, Suite 300
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: pc-tos/ms-dos
SOFTWARE: Patentia Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07789,913
FILING DATE: 19911112
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07-07-761,766
FILING DATE: 02-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/440,094
FILING DATE: 22-NOV-1989
ATTORNEY/AGENT INFORMATION:
NAME: Stratford, Carol A.
REGISTRATION NUMBER: 34,444
REFERENCE/AGENT NUMBER: 5865-0005,30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0860
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 1 amino acids
TYPE: AMINO ACID

TOPOLOGY: both
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: peptide fragment used in the claims
US-07-789-913-23

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 C 1

RESULT 6
US-07-789-913-25
Sequence 25, Application US/07789913
Patent No. 555095
GENERAL INFORMATION:
APPLICANT: Miljanich, George P.
APPLICANT: Homeros, Stephen S.
APPLICANT: Fox, James A.
APPLICANT: Valentino, Karen L.
APPLICANT: Bittner, Robert S.
APPLICANT: Yamashiro, Donald H.
TITLE OF INVENTION: Delayed Treatment Method of Reducing
TITLE OF INVENTION: Ischemia Related Neuronal Damage
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Peter Dehlinger
STREET: 450 Cambridge Avenue, Suite 300
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: pc-tos/ms-dos
SOFTWARE: Patentia Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07789,913
FILING DATE: 19911112
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/561,766
FILING DATE: 02-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/440,094
FILING DATE: 22-NOV-1989
ATTORNEY/AGENT INFORMATION:
NAME: Stratford, Carol A.
REGISTRATION NUMBER: 34,444
FILING DATE: 02-AUG-1990
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0860
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 1 amino acids
TYPE: AMINO ACID
MOLECULE TYPE: both
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: peptide fragment used in the claims
US-07-789-913-25

```

Query Match          0.0% Score 0: DB 1: Length 1:
Best Local Similarity 0.0% Prod. No. 0:
Matches 0: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

QY 1 X 1
DB 1 R 1

RESULT 7
US-08-049-794-25
Sequence 25, Application US/08049794
Patent No. 5587454
GENERAL INFORMATION:
APPLICANT: JUSTICE, ALAN
APPLICANT: SINGH, PETERINDER
APPLICANT: GOHIL, KISHOR C
APPLICANT: VALENTINO, KAREN L
APPLICANT: MILANICH, GEORGE P
TITLE OF INVENTION: METHODS OF PREVENTING ANAESTHESIA AND
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Peter Deblinger
STREET: 450 Cambridge Avenue, Suite 300
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIA TYPE: floppy disk
COMPILED: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 05/08049794
FILING DATE: 19930415
ATTORNEY/AGENT INFORMATION:
NAME: Stratford, Carol A.
REGISTRATION NUMBER: 34,444
REFERENCE/AGENT NUMBER: 5965-0093, 00
TELEPHONE: (415) 824-0860
TELEFAX: (415) 824-0960
INFORMATION FOR SEQ TO NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 1 amino acids
MOLECULE TYPE: linear
TOPLOGY: linear
HYDROPHOBIC: no
INDIVIDUAL ISOLATED: COMPETITIVE GROUP 1 FRAGMENT, PAGE
US-08-049-794-25

Query Match          0.0% Score 0: DB 1: Length 1:
Best Local Similarity 0.0% Prod. No. 0:
Matches 0: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

QY 1 X 1
DB 1 R 1

RESULT 8
US-08-049-794-25

```

```

Sequence 25, Application US/08049794
Patent No. 5587454
GENERAL INFORMATION:
APPLICANT: JUSTICE, ALAN
APPLICANT: SINGH, PETERINDER
APPLICANT: GOHIL, KISHOR C
APPLICANT: VALENTINO, KAREN L
APPLICANT: MILANICH, GEORGE P
TITLE OF INVENTION: METHODS OF PREVENTING ANAESTHESIA AND
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Peter Deblinger
STREET: 450 Cambridge Avenue, Suite 300
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIA TYPE: floppy disk
COMPILED: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 05/08049794
FILING DATE: 19930415
ATTORNEY/AGENT INFORMATION:
NAME: Stratford, Carol A.
REGISTRATION NUMBER: 34,444
REFERENCE/AGENT NUMBER: 5965-0093, 00
TELEPHONE: (415) 824-0860
TELEFAX: (415) 824-0960
INFORMATION FOR SEQ TO NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 1 amino acids
MOLECULE TYPE: linear
TOPLOGY: linear
HYDROPHOBIC: no
INDIVIDUAL ISOLATED: COMPETITIVE GROUP 1 FRAGMENT, PAGE
US-08-049-794-25

Query Match          0.0% Score 0: DB 1: Length 1:
Best Local Similarity 0.0% Prod. No. 0:
Matches 0: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

QY 1 X 1
DB 1 R 1

RESULT 9
US-08-433-037-12
Sequence 12, Application US/08433017
Patent No. 5707828
GENERAL INFORMATION:
APPLICANT: Steekrishna, Kotikanyadan
APPLICANT: Bart, Kathryn A.
APPLICANT: Hirtley, Russell A.
APPLICANT: Thill, Gregory P.
APPLICANT: Ischopp, Jerald F.
TITLE OF INVENTION: EXPRESSION OF HUMAN SENSORY ALBUMIN IN
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:

```

```

ADDRESS: Seilly, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: U.S.A.
ZIP: 11540-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/66/443,677
FILING DATE: 03-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: DIGITAL, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/WORKET NUMBER: 9108Z
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-1125
TELEX: 240 901 SANS UR
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-448-606-12

```

```

Query Match: 0.0%, Score 0; DB 1; Length 1;
Host Local Similarity: 0.0%; Freq. No. 0;
Matches: 0; Conservative: 0; Mismatches: 1; Indels: 0; Gaps: 0;

```

```

UY 1 X 1
DB 1 1 1

```

```

RESULT 10
US-08-448-606-4
Sequence 4, Application US/08448606
Patent No. 572114
GENERAL INFORMATION:
APPLICANT: Abrahams D, Lars
APPLICANT: Kaldor D, Christina
APPLICANT: Lake, Mats
APPLICANT: Mikalson, SA
ABSTRACT: Seilly, Scott, Murphy & Presser
TITLE OF INVENTION: Peptide System For Protein
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESS: Seilly, Scott, Murphy & Presser
STREET: 1900 M Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: U.S.
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/448,606
FILING DATE: 25-AUG-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01061

```

```

FILING DATE: 03-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9203754-0
FILING DATE: 11-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Amelick, Burton A.
REGISTRATION NUMBER: 24,852
REFERENCE/WORKET NUMBER: 0151/20121
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 331-7111
TELEFAX: (202) 293-6259
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1 amino acid
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-448-606-4

```

```

Query Match: 0.0%, Score 0; DB 1; Length 1;
Host Local Similarity: 0.0%; Freq. No. 0;
Matches: 0; Conservative: 0; Mismatches: 1; Indels: 0; Gaps: 0;
UY 1 X 1
DB 1 0 1

```

```

RESULT 11
US-07-869-933-16
Sequence 16, Application US/07869933
Patent No. 5770396
GENERAL INFORMATION:
APPLICANT: Kinet, Jean-Pierre
TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESS: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/869,933
FILING DATE: 19920416
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Behr, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/WORKET NUMBER: 48449/154 NIND
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 1 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-869-933-16

```

```

Query Match: 0.0%, Score 0; DB 1; Length 1;

```

Best Local Similarity: 0.0%; Pred. No. 0;
Matches: 0; Conservative: 0; Mismatches: 1; Indels: 0; Gaps: 0;

QY 1 X 1
14 1 M 1

RESULT 12
US-08-294-150A-24
Sequence 24, Application US/09/294150A
Patent No. 5792629
GENERAL INFORMATION:
APPLICANT: MORISHITA, HIROAKI
APPLICANT: KAMAMORI, TOSHIOKI
APPLICANT: NISHIKAWA, MASAHITO
TITLE OF INVENTION: PEPTIDES, DNA FRAGMENT ENCODING THE
TITLE OF INVENTION: SAME AND PROCESS FOR PREPARING THE SAME, AND ENZYME
TITLE OF INVENTION: INHIBITION PROCESS, TISSUE COMPOSITION AND METHODS OF
TITLE OF INVENTION: TREATING USING THE SAME
NUMBER OF SEQUENCES: 110
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DWANE, SWICKER & MATHIS
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIA TYPE: floppy disk
MODIFIER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 03/797931, USA
FILING DATE: 19-AUG-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 02/677911, 213
FILING DATE: 13-NOV-1990
PRELIMINARY INVENTION DATA:
PRELIMINARY INVENTION DATA:
APPLICATION NUMBER: JP 2-00745
FILING DATE: 13-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Morishita, Hiroaki M.
REGISTRATION NUMBER: 46,607
REFERENCE/AGENT NUMBER: 020650 049
TELEPHONE: (703) 846-6620
TELEFAX: (703) 846-2021
INVENTION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 1 amino acids
TYPE: amino acid
Topology: linear
MOLECULE TYPE: peptide
US-08-294-150A-24

Query Match: 0.0%; Score 0; DR 1; Length 1;
Best Local Similarity: 0.0%; Pred. No. 0;
Matches: 0; Conservative: 0; Mismatches: 1; Indels: 0; Gaps: 0;
QY 1 X 1
14 1 C 1
RESULT 13
US-08-294-150A-40
Sequence 40, Application US/09/294150A
Patent No. 5792629
GENERAL INFORMATION:

APPLICANT: MORISHITA, HIROAKI
APPLICANT: KAMAMORI, TOSHIOKI
APPLICANT: NISHIKAWA, MASAHITO
TITLE OF INVENTION: PEPTIDES, DNA FRAGMENT ENCODING THE
TITLE OF INVENTION: SAME AND PROCESS FOR PREPARING THE SAME, AND ENZYME
TITLE OF INVENTION: INHIBITION PROCESS, TISSUE COMPOSITION AND METHODS OF
TITLE OF INVENTION: TREATING USING THE SAME
NUMBER OF SEQUENCES: 110
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DWANE, SWICKER & MATHIS
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIA TYPE: floppy disk
MODIFIER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 03/797931, USA
FILING DATE: 19-AUG-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/791,213
FILING DATE: 13-NOV-1990
PRELIMINARY INVENTION DATA:
APPLICATION NUMBER: JP 2-00745
FILING DATE: 13-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Morishita, Hiroaki M.
REGISTRATION NUMBER: 46,607
REFERENCE/AGENT NUMBER: 020650-049
TELEPHONE: (703) 846-6620
TELEFAX: (703) 846-2021
INVENTION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 1 amino acids
TYPE: amino acid
Topology: linear
MOLECULE TYPE: peptide
US-08-294-150A-40

Query Match: 0.0%; Score 0; DR 1; Length 1;
Best Local Similarity: 0.0%; Pred. No. 0;
Matches: 0; Conservative: 0; Mismatches: 1; Indels: 0; Gaps: 0;

QY 1 X 1
Db 1 C 1

RESULT 14
US-08-496-847-23
Sequence 23, Application US/08496847
Patent No. 5795864
GENERAL INFORMATION:
APPLICANT: AMSTUTZ, GARY A.
APPLICANT: BOWSTON, STEPHEN S.
APPLICANT: GOHIL, KISHORCHANDRA
APPLICANT: ADJAINOUS, PETER L.
APPLICANT: KRISHNAPATI, RAMASHARMA
TITLE OF INVENTION: METHODS AND
TITLE OF INVENTION: FORMULATIONS FOR PREVENTING PROGRESSION OF PREDIABETIC PAIN
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dellinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA

COUNTRY: US
 ZIP: 94306-1546
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: PASTSEO for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: 9509496,847
 FILING DATE: 27-JUN-1995
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Stratford, Carol A
 REGISTRATION NUMBER: 34,444
 REFERENCE/CHECK NUMBER: 5925 0009,31
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-324-0880
 TELEFAX: 650-324-0960
 INFORMATION FOR SEQ ID NO: 25:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: CONOPEPTIDE GROUP 1 FRAGMENT
 US-08-496-847-25

Query Match 0.0% Score 0; DB 1; Length 1;
 Host local Similarity 0.0% Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 DB 1 C 1

RESULT 15
 US-08-496-847-25
 Sequence 25, Application US/08496847
 Patient No. 5795864
 GENERAL INFORMATION:
 APPLICANT: Amstutz, Gary A.
 APPLICANT: Bokorox, Stephen S.
 APPLICANT: Gohil, Kishorhantra
 APPLICANT: Ahlenssens, Peter I.
 APPLICANT: Kristipati, Ramasharma
 TITLE OF INVENTION: METHODS AND
 NUMBER OF STUDENTS: 46
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dehlinger & Associates
 STREET: 450 Cambridge Avenue, Suite 250
 CITY: Palo Alto
 STATE: CA
 COUNTRY: US
 ZIP: 94306-1546
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: PASTSEO for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: 9509496,847
 FILING DATE: 27-JUN-1995
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Stratford, Carol A
 REGISTRATION NUMBER: 34,444
 REFERENCE/CHECK NUMBER: 5925 0009,31
 TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-324-0880
 TELEFAX: 650-324-0960
 INFORMATION FOR SEQ ID NO: 25:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: CONOPEPTIDE GROUP 1 FRAGMENT
 US-08-496-847-25

Query Match 0.0% Score 0; DB 1; Length 1;
 Host local Similarity 0.0% Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 X 1
 DB 1 R 1

Search completed: September 30, 2001, 10:29:48
 Job time: 48 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd

• M protein - protein search, using sw model

Received: September 40, 2001, 10:40:46 ; Search time: 32.59 seconds

(microliters)
16.234 mg/ml \times 100 μ l = 1.6234 g

Title:	US-00-44
Percent Score:	4
Sequence:	1 XXXX 4

Scoring table

Figure 10: $\log_{10}(\text{variance})$ vs. $\log_{10}(\text{mean})$

Notched: 425026 90911, 132295027 100110000

Total number of hits satisfying chosen parameters: 425026

Author	Year	Country	Sample Size	Study Design	Findings
Smith et al.	2001	USA	1,200	Longitudinal	Increased risk of depression in children of parents with mental illness.
Johnson et al.	2003	UK	800	Cross-sectional	Family history of mental illness associated with higher levels of anxiety.
Lee et al.	2005	Canada	1,500	Longitudinal	Genetic factors play a significant role in the development of bipolar disorder.
Wong et al.	2007	Australia	900	Cross-sectional	Environmental factors such as childhood trauma can trigger mental health issues.
Chen et al.	2009	China	2,100	Longitudinal	Family environment significantly influences the onset of schizophrenia.
Miller et al.	2011	USA	1,800	Cross-sectional	Stressful life events can exacerbate symptoms of major depressive disorder.
Nguyen et al.	2013	Vietnam	1,100	Longitudinal	War-related trauma has long-term effects on mental health in Vietnam.
Patel et al.	2015	India	1,300	Cross-sectional	Cultural beliefs and family support influence the management of mental illness.
Kim et al.	2017	South Korea	1,600	Longitudinal	Genetic predisposition interacts with environmental factors in the development of autism.
Roberts et al.	2019	USA	1,400	Cross-sectional	Family history of substance use disorder is linked to higher rates of addiction.
Yamamoto et al.	2021	Japan	1,700	Longitudinal	Social isolation and family structure impact the prevalence of dementia.

Maximum LH seq length: 20000000000

Post-processing: Minimum Match Offset

List in list 45 summaries

[illegible]

```

1:  SP_HMM1: *
2:  SP_AutoTic: *
3:  SP_AutoTic: *
4:  SP_HMM1: *
5:  SP_HMM1: *
6:  SP_InvertAuto: *
7:  SP_InvertAuto: *
8:  SP_InvertAuto: *
9:  SP_Phase: *
10: SP_Phase: *
11: SP_Phase: *
12: SP_Phase: *
13: SP_Phase: *
14: SP_Phase: *

```

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is defined by analysis of the total score distribution.

SUMMARY

Result	No.	Score	Match	Length	DB	ID	Insertion
1	0	0.0	4	11	Q08442	Q08442	catenulatus
2	0	0.0	5	10	Q09007	Q09007	hordeum
3	0	0.0	5	13	P82070	P82070	illoria
4	0	0.0	5	13	P82071	P82071	illoria
5	0	0.0	5	13	P82072	P82072	illoria
6	0	0.0	5	13	P82073	P82073	illoria
7	0	0.0	5	13	P82099	P82099	illoria
8	0	0.0	5	13	P82100	P82100	illoria
9	0	0.0	6	2	P82455	P82455	lactobacilli
10	0	0.0	6	14	Q08720	Q08720	homo sapiens
11	0	0.0	6	10	P82181	P82181	spinacia
12	0	0.0	6	10	P82182	P82182	spinacia
13	0	0.0	6	10	P82541	P82541	spinacia
14	0	0.0	6	13	P82066	P82066	illoria
15	0	0.0	7	2	Q05448	Q05448	strepomyces
16	0	0.0	7	2	Q07454	Q07454	strepomyces
17	0	0.0	7	2	Q43028	Q43028	pseudomonas
18	0	0.0	7	2	Q47029	Q47029	enterobacter
19	0	0.0	7	2	P70804	P70804	actinobacter

ALIGNMENT

20	0	0.0	7	2	P22591	<i>actinobacil</i>
21	0	0.0	7	2	O50556	<i>actinobacil</i>
22	0	0.0	7	2	O47505	<i>escherichia</i>
23	0	0.0	7	2	O47477	<i>escherichia</i>
24	0	0.0	7	4	O15897	<i>homo sapien</i>
25	0	0.0	7	4	O15900	<i>homo sapien</i>
26	0	0.0	7	6	O28742	<i>cytolacrus</i>
27	0	0.0	7	8	P92421	<i>psathyrosta</i>
28	0	0.0	7	8	P92385	<i>psathyrosta</i>
29	0	0.0	7	8	P52210	<i>hardum mar</i>
30	0	0.0	7	8	P92214	<i>hardum c</i>
31	0	0.0	7	8	P92218	<i>anthracopyr</i>
32	0	0.0	7	8	P92221	<i>throms ther</i>
33	0	0.0	7	8	P92226	<i>ci lihopsts</i>
34	0	0.0	7	8	P92372	<i>hynaldia v</i>
35	0	0.0	7	8	P92381	<i>hardum bra</i>
36	0	0.0	7	8	P92387	<i>hardum p</i>
37	0	0.0	7	8	P92390	<i>hardum p</i>
38	0	0.0	7	8	P92393	<i>hardum vti</i>
39	0	0.0	7	8	P92425	<i>hardum vti</i>
40	0	0.0	7	8	P92427	<i>hardum vti</i>
41	0	0.0	7	8	P92430	<i>hardum vti</i>
42	0	0.0	7	8	P92442	<i>hardum vti</i>
43	0	0.0	7	8	P92440	<i>hardum vti</i>
44	0	0.0	7	8	O95945	<i>throms ther</i>
45	0	0.0	7	10	P93233	<i>throms ther</i>

So submit: 4 AA, 475 MW, 633732-2420000000 - TC643

```

Query Match      0.08; Score 0; Bk 11; Length 4;
Host Local Similarity 0.08; Pred. No. 0;
Matches of Conservative 0; Mismatches 1; Indels 0; Gaps 0

```

γ	1	8	1
δ	1	2	1

"The structure of new peptides from the Australian red tree-fern (*Adiantum reginae*), the skin peptide, is given as a probe for the study of evolutionary trends of amphibians." *Aust. J. Chem.* 49:195, 964 (1996).

ANTHROPIC ACTIVITY

OT -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
 KW -1- MASS SPECTROMETRY: MW 655; METH-0 FAB.
 FT Amphibian skin: Amidation.
 FT MOD-RES 5 AA: 5 AMITATION.
 SO SEQUENCE 5 AA: 556 MW: 71A9C9C810300000 C8C64;

Query Match 0.0%; Score 0; DB 13; Length 5;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservation 0; Mismatches 1; Indels 0; Gaps 0;

UY 1 X 1
 DB 1 1 1

RESULT 6
 ID P82073 PRELIMINARY: PRI: 5 AA.
 AC P82073;
 DT 01-MAY-2000 (TREMUR: 13, Created)
 DT 01-MAY-2000 (TREMUR: 13, last sequence update)
 DT 01-MAY-2000 (TREMUR: 13, last annotation update)
 DE RUBELLIN 3.2.
 OS Litoria rubella (Desert tree frog).
 SC Amphibia: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
 SC Amphibia: Batrachia: Anura: Neobatrachia: Hylodidae: Hylidae:
 SC Litoria.
 SC Nhl_TaxID: 104895;
 RN 11
 RP SEQUENCE.
 RC TISSUE SKIN SECRETION.
 EA Wabnitz P.A., Bowle J.H., Tyler M.J., Wallace J.C.
 FT "Peptides from the skin glands of the Australian Popping tree frog
 Litoria electrica. Comparison with the skin peptides from Litoria
 rubella."
 RL Aust. J. Chem. 52:0-0(1999).
 CC -1- FUNCTION: CAPRIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
 CC ANTHROPIC ACTIVITY.
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
 KW Amphibian skin.
 SO SEQUENCE 5 AA: 570 MW: 71A9C9C82A000000 C8C64;

Query Match 0.0%; Score 0; DB 13; Length 5;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservation 0; Mismatches 1; Indels 0; Gaps 0;

UY 1 X 1
 DB 1 1 1

RESULT 7
 ID P82094 PRELIMINARY: PRI: 5 AA.
 AC P82094;
 DT 01-MAY-2000 (TREMUR: 13, Created)
 DT 01-MAY-2000 (TREMUR: 13, last sequence update)
 DT 01-MAY-2000 (TREMUR: 13, last annotation update)
 DE ELECTRIN 4.
 OS Litoria rubella (Desert tree frog).
 SC Amphibia: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
 SC Amphibia: Batrachia: Anura: Neobatrachia: Hylodidae: Hylidae:
 SC Litoria.
 SC Nhl_TaxID: 104895;
 RN 11
 RP SEQUENCE.
 RC TISSUE SKIN SECRETION.
 EA Wabnitz P.A., Bowle J.H., Tyler M.J., Wallace J.C.
 FT "Peptides from the skin glands of the Australian Popping tree frog
 Litoria electrica. Comparison with the skin peptides from Litoria

RT rubella."
 RL Aust. J. Chem. 52:0-0(1999).
 KW Amphibian skin: Amidation.
 FT MOD-RES 5 AA: 5 AMITATION.
 SO SEQUENCE 5 AA: 650 MW: 66A761F2C9A00000 C8C64;

Query Match 0.0%; Score 0; DB 13; Length 5;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservation 0; Mismatches 1; Indels 0; Gaps 0;

UY 1 X 1
 DB 1 1 1

RESULT 8
 ID P82100 PRELIMINARY: PRI: 5 AA.
 AC P82100;
 DT 01-MAY-2000 (TREMUR: 13, Created)
 DT 01-MAY-2000 (TREMUR: 13, last sequence update)
 DT 01-MAY-2000 (TREMUR: 13, last annotation update)
 DE ELECTRIN 4.
 OS Litoria rubella (Desert tree frog).
 SC Amphibia: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
 SC Amphibia: Batrachia: Anura: Neobatrachia: Hylodidae: Hylidae:
 SC Litoria.
 SC Nhl_TaxID: 104895;
 RN 11
 RP SEQUENCE.
 RC TISSUE SKIN SECRETION.
 EA Wabnitz P.A., Bowle J.H., Tyler M.J., Wallace J.C.
 FT "Peptides from the skin glands of the Australian Popping tree frog
 Litoria electrica. Comparison with the skin peptides from Litoria
 rubella."
 RL Aust. J. Chem. 52:0-0(1999).
 CC -1- FUNCTION: CAPRIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
 CC ANTHROPIC ACTIVITY.
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
 KW Amphibian skin: Amidation.
 SO SEQUENCE 5 AA: 616 MW: 61F201A05A000000 C8C64;

Query Match 0.0%; Score 0; DB 13; Length 5;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservation 0; Mismatches 1; Indels 0; Gaps 0;

UY 1 X 1
 DB 1 1 1

RESULT 9
 ID P82655 PRELIMINARY: PRI: 6 AA.
 AC P82655;
 DT 01-MAR-2001 (TREMUR: 16, Created)
 DT 01-MAR-2001 (TREMUR: 16, last sequence update)
 DT 01-MAR-2001 (TREMUR: 16, last annotation update)
 DE ACID SHOCK PROTEIN (FRAGMENT).
 OS Lactobacillus sanfranciscensis (Lactobacillus sanfranciscensis).
 SC Bacteria: Firmicutes: Bacillus/Clostridium group: Lactobacillaceae:
 SC Lactobacillus.
 SC Nhl_TaxID: 1625;
 RN 11
 RP SEQUENCE.
 RC STRAIN CBI.
 EA De Angelis M., Rini L., Pallini V., Gobbelli M.
 FT "The acid stress response in Lactobacillus sanfranciscensis."
 CC Microbiology 0:0-0(2000).
 CC 1 INDUCTION: OVEREXPRESSED IN ACID ENVIRONMENTS.
 FT MOD-RES 6 AA: 6
 SO SEQUENCE 6 AA: 778 MW: 6A4F5E8F72A50000 1E564;

Query Match: 0.0%; Score 0; DB 2; Length 6;
 Post Local Similarity: 100.0%; Pred. No. 0;
 Matches: 1; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

QY 1 X 1
 DB 1 X 1

RESULT 10
 GDB20
 ID: 008720 PRELIMINARY: PRT: 6 AA.
 AC: 008720;
 DT: 01-NOV-1996 (TREMBL: 01, last sequence update)
 DI: 01-NOV-1996 (TREMBL: 01, last annotation update)
 DE: Y PROTEIN (FRAMMENT).
 GN: CREB.
 OS: Homo sapiens (human).
 OC: Eukaryota; Metazoa; Chordata; Mammalia; Vertebrata; Euteleostomi;
 NC: Mammalia; Eutheria; Primates; Catartida; Homalidae; Homo.
 NX: NPI_taxid: 9606.
 LN: 11
 RF: SEQUENCE FROM N.A.
 RA: MEDLINE 9010601; PubMed 196444;
 RW: Weber G., Hedberg J.E.;
 RT: "Novel testis germ cell-specific transcript of the 'B6' gene encodes an alternatively spliced exon with multiple in-frame stop codons."
 RI: Endocrinology 141:2010-2015(1992).
 DR: EMBL: X60994; CAA8780.1; .
 FI: N_N_TER 1
 FT: N_N_TER 5
 SO: SEQUENCE 6 AA: 675 MW: 67228367; EMBL: X6094.

Query Match: 0.0%; Score 0; DB 4; Length 6;
 Post Local Similarity: 0.0%; Pred. No. 0;
 Matches: 0; Conservative: 0; Mismatches: 1; Indels: 0; Gaps: 0;

QY 1 X 1
 DB 1 S 1

RESULT 11
 PRELIMINARY: PRT: 6 AA.
 ID: P82181;
 AC: P82181;
 DT: 01-JUN-2000 (TREMBL: 14, last sequence update)
 DI: 01-JUN-2000 (TREMBL: 14, last annotation update)
 DE: 01-MAR-2001 (TREMBL: 16, last annotation update)
 DE: CHLOROPLAST 50S RIBOSOMAL PROTEIN L10 BETA (FRAGMENT).
 OS: Spirochaetaceae (Spirocha).
 NC: Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC: Magnoliophyta; eudicotyledons; core eudots; Caryophyllales;
 NX: NPI_taxid: 5562;
 LN: 11
 RF: SEQUENCE.
 RA: STRAIN CV. ALMABO; TISSUE LEAF;
 RW: MEDLINE 2043798; PubMed 10874046;
 RT: Yamaguchi K., Subramanian A.R.;
 RT: "The plastid ribosomal proteins. Identification of all the proteins in the 50S subunit of an orange-red ribosome (chloroplast)."
 CC: -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
 CC: -1- SUBCELLULAR LOCATION: CHLOROPLAST.
 CC: -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
 CC: -1- MISCELLANEOUS: ON THE 23S RIB. ITS MW IS: 16.5 KDA.
 CC: -1- SIMILARITY: EMBL:R08 T- THE L10F FAMILY OF RIBOSOMAL PROTEINS.
 DR: InterPro: IPR002463;
 DR: PROSITE: PS01109; RIBOSOMAL_L10F PARTIAL.
 KW: Ribosomal protein; chloroplast; rRNA-binding.
 FT: N_N_TER 6
 FT: N_N_TER 6
 SO: SEQUENCE 6 AA: 675 MW: 64218415; EMBL: X6094.

KW Ribosomal protein; chloroplast; rRNA-binding.
 FT N_N_TER 6
 SO SEQUENCE 6 AA: 675 MW: 64218415; EMBL: X6094.

Query Match: 0.0%; Score 0; DB 10; Length 6;
 Post Local Similarity: 0.0%; Pred. No. 0;
 Matches: 0; Conservative: 0; Mismatches: 1; Indels: 0; Gaps: 0;

QY 1 X 1
 DB 1 A 1

RESULT 12
 P82182
 ID: P82182 PRELIMINARY: PRT: 6 AA.
 AC: P82182;
 DT: 01-JUN-2000 (TREMBL: 14, last sequence update)
 DI: 01-JUN-2000 (TREMBL: 14, last annotation update)
 DE: 01-MAR-2001 (TREMBL: 16, last annotation update)
 DE: CHLOROPLAST 50S RIBOSOMAL PROTEIN L10 GAMMA (FRAGMENT).
 OS: Spirochaetaceae (Spirocha).
 NC: Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC: Magnoliophyta; eudicotyledons; core eudots; Caryophyllales;
 NX: NPI_taxid: 5562;
 LN: 11
 RF: SEQUENCE.
 RA: STRAIN CV. ALMABO; TISSUE LEAF;
 RW: MEDLINE 2043798; PubMed 10874046;
 RT: Yamaguchi K., Subramanian A.R.;
 RT: "The plastid ribosomal proteins. Identification of all the proteins in the 50S subunit of an orange-red ribosome (chloroplast)."
 CC: -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
 CC: -1- SUBCELLULAR LOCATION: CHLOROPLAST.
 CC: -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
 CC: -1- MISCELLANEOUS: ON THE 23S RIB. ITS MW IS: 16.5 KDA.
 CC: -1- SIMILARITY: EMBL:R08 T- THE L10F FAMILY OF RIBOSOMAL PROTEINS.
 DR: InterPro: IPR002463;
 DR: PROSITE: PS01109; RIBOSOMAL_L10F PARTIAL.
 KW: Ribosomal protein; chloroplast; rRNA-binding.
 FT: N_N_TER 6
 FT: N_N_TER 6
 SO: SEQUENCE 6 AA: 675 MW: 64218415; EMBL: X6094.

Query Match: 0.0%; Score 0; DB 10; Length 6;
 Post Local Similarity: 0.0%; Pred. No. 0;
 Matches: 0; Conservative: 0; Mismatches: 1; Indels: 0; Gaps: 0;

QY 1 X 1
 DB 1 A 1

RESULT 13
 P82541
 ID: P82541 PRELIMINARY: PRT: 6 AA.
 AC: P82541;
 DT: 01-OCT-2000 (TREMBL: 15, last sequence update)
 DI: 01-OCT-2000 (TREMBL: 15, last sequence update)
 DE: 01-MAR-2001 (TREMBL: 16, last annotation update)
 DE: CHLOROPLAST 50S RIBOSOMAL PROTEIN S15 BETA (FRAMMENT).
 OS: Spirochaetaceae (Spirocha).
 NC: Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC: Magnoliophyta; eudicotyledons; core eudots; Caryophyllales;
 NX: NPI_taxid: 5562;
 LN: 11
 RF: SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
 RA: STRAIN CV. ALMABO; TISSUE LEAF;

```

RX MEDLINE-20435797; PubMed 10874039;
KA Yamanechi K., von Knorrich K., Subramanian A.B.;
KI The small ribosomal proteins. Identification of all the proteins in
KL the small subunit of an organelle ribosome (chloroplast)";
RL J. Biol. Chem. 273:28455-28465(2000);
OT FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 16S RIBOSOMAL RNA.
OO SEQUENCE LOCATION: CHLOROPLAST.
OT LIGASE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
OT MASS SPECTROMETRY: MW:10477.0; METHOD ELECTROSPRAY.
OT MASS SEQUENCE: MW:10470; METHOD ELECTROSPRAY.
OT MISCELLANEOUS: SIX ALPHA AND BETA FORMS DIFFER IN THE SIX BETA
OT FORM IS THE MINOR BASIC FORM.
OT MISCELLANEOUS: ON THE 2D-GEL, THE MW IS 12 KDA.
OT SIMILARITY: BELONGS TO THE S19 FAMILY OF RIBOSOMAL PROTEINS.
DK INTERPRO: IPR009222;
DK PROSITE: PS00922; RIBOSOMAL_S19; PARTIAL;
DK RIBOSOMAL PROTEIN; CHLOROPLAST; RNA-BINDING.
FI N_N_TERM 6;
SI SEQUENCE 6 AA: 792 MW: 6333376A11C006 CREC64;

Query Match 0.08; Score 0; DB 10; Length 6;
Best local Similarity 0.08; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 T 1

RESULT 14
PRT 36
II 192096; PRELIMINARY; PRT: 6 AA.
DI 01-MAY-2000 (TREMUR-01, Created)
DI 01-MAY-2000 (TREMUR-01, Last sequence update)
DI 01-MAY-2000 (TREMUR-01, Last annotation update)
DE ELEKTRIN 1.
OS Litoria rubella (Poison tree frog).
OC Eukaryota; Metazoa; Chordata; Grenata; Vertebrata; Entolestomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria;
OC X Nbr_LaxID: 104895;
KN 11;
RP SEQUENCE.
RP TISSUE:SKIN SECRETION;
KA Wainitz P.A., Howie J.H., Tyler M.J., Wallace J.C.;
KI "Toxins from the skin glands of the Australian buzzy tree frog
KI Litoria electrica: comparison with the skin peptides from Litoria
KI rubella";
RL Aust. J. Chem. 52:0-0(1999).
KW Amphibian skin; Amudation.
FI M_LRES 6;
SI SEQUENCE 6 AA: 792 MW: 65837647729A000 CREC64;

Query Match 0.08; Score 0; DB 13; Length 6;
Best local Similarity 0.08; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 F 1

RESULT 15
C54248
II 054248; PRELIMINARY; PRT: 7 AA.
DI 01-NOV-1996 (TREMUR-01, Created)
DI 01-NOV-1996 (TREMUR-01, Last sequence update)
DI 01-NOV-1996 (TREMUR-01, Last annotation update)
DE SPOY & ADK GINES (FRAGMENT).

```

```

GN RFL1.
OS Streptomyces griseus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomyces; Streptomyces.
OX NBR_LaxID:1911;
RN 11;
RP SEQUENCE FROM N.A.
KC STRAIN-N2-3-11;
KA Pochlins S., Petersberg W., Wehner U.F.;
FI Submitted (Feb 1996) to the EMBL/GenBank/DDBJ databases.
DK EMBL: X95915; CAA05160.1;
FI N_N_TERM 1;
SI SEQUENCE 7 AA: 760 MW: 7207801B01B2A0 CREC64;

```

```

Query Match 0.08; Score 0; DB 2; Length 7;
Best local Similarity 0.08; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 V 1

```

Search completed: September 30, 2001, 10:30:47
 Job time: 117 sec

P1 Bertram PM, Gansseman YOL, Jespers LS, Laroche YR,
 P1 Lauwereys MJ, Messens JHL, Moyle M, Stanssens PPH,
 P1 Vlasuk GP;
 XX
 XX WPI: 1996-220037/22
 DR N-PSDB; AAT12947.
 XX
 PT Proteolysis with anticoagulant and/or serine protease inhibitory
 P1 activity - isolated from nematodes and useful to inhibit blood
 P1 coagulation
 XX
 PS Claim 89 + 96, Fig 9, 24pp, English.
 CC
 CC Proteins with anticoagulant and/or serine protease inhibitory
 CC activity, isolated from nematodes, are useful to inhibit blood
 CC coagulation. The proteins can be added to blood collection tubes
 CC defining the collection of mammalian plasma. They are also useful
 CC to prevent or inhibit thrombosis, and may be given alone or in
 CC combination with other therapeutic or in vivo diagnostic agents.
 CC The proteins can serve as immunogens to raise antibodies for use in
 CC the diagnosis and identification of NAP concn. levels in biological
 CC fluids, e.g. to detect mammalian infection with a parasitic worm.
 CC They can also be used as immunogens in prophylactic and therapeutic
 CC vaccines against parasitic worm infection. The proteins may
 CC double the clotting time of human plasma in prothrombin time assays
 CC when present at 10-50 nmol, and double the clotting time of human
 CC plasma in activated partial thrombin time assays when present
 CC at 10-100 nmol.
 CC The anticoagulant proteins are pref. derived from
 CC *Ancylostoma caninum*, *A. ceylanicum*, *A. dirdenale*, *Necator*
 CC *americanus* or *Heligmosomoides polygyrus*.
 CC The proteins first have 2 NAP domains and specifically inhibit
 CC the catalytic activity of the factor VIIa/TF complex in the
 CC presence of factor Xa or a catalytically inactive factor Xa deriv.,
 CC do not specifically inhibit the activation of factor VIIa in the
 CC absence of TF and do not specifically inhibit prothrombinase
 CC
 CC Sequence 91 AA.
 SO
 Query Match 100.0%; Score 486; DB 17; Length 91;
 Host Local Similarity 100.0%; Pred. No. 66-39;
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KATMGCGENFYDSCGSKETKTKYKGVPEFPIHFNVPCLVWVHGVVPEGFRNK 50
 Lb 8 KATMGCGENFYDSCGSKETKTKYKGVPEFPIHFNVPCLVWVHGVVPEGFRNK 67
 QY 61 DDKCVSAEDCELDMMDFYPQTRN 84
 DB 68 ddkcvsaedceldmmdflypqrtn 91
 RESULT 4
 AAY30393
 ID AAY30393 standard; Protein; 91 AA.
 AC AAY30393;
 XX
 XX 15-NOV-1999 (first entry)
 DE Nematode extracted anticoagulant protein AcanAPc2.
 AC
 AC Nematode extracted anticoagulant protein; NAP; anticoagulant;
 KW serine protease inhibitor; NAP domain; factor VIIa/TF.
 XX
 XX *Ancylostoma caninum*.
 OS
 XX US955294-A.
 FN
 XX 21-SEP-1999.
 PD
 XX 19-APR-1996; 9605-0634641.
 PF

XX
 XX 19-APR-1996; 9605-0634641.
 PF 18-OCT-1994; 9405-0326110.
 PR 05-JUN-1995; 9505-0461965.
 PR 05-JUN-1995; 9505-0465980.
 PR 05-JUN-1995; 9505-0486397.
 PR 05-JUN-1995; 9505-0486399.
 PR 17-OCT-1995; 9505-051431.
 XX
 XX (CORV-) CORVAS INT INC.
 XX
 P1 Bertram PM, Gansseman YOL, Jespers LS, Laroche YR,
 P1 Lauwereys MJ, Messens JHL, Moyle M, Stanssens PPH,
 P1 Vlasuk GP;
 XX
 XX WPI: 1996-220037/22
 DR N-PSDB; AAT29999.
 XX
 PT Screening an isolated protein for Nematode extracted Anticoagulant
 P1 protein domains
 XX
 XX Example 13; Fig 9; 197pp; English.
 CC The present sequence represents a nematode extracted anticoagulant
 CC protein (NAP). The protein has activity as an anticoagulant and/or serine
 CC protease inhibitor. The protein contains at least one NAP domain which
 CC has selective inhibitory activity for factor VIIa/TF. The specification
 CC describes a method for screening an isolated protein at least one domain
 CC for factor VIIa/TF selective inhibitory activity. The method comprises
 CC determining the time to clotting effected by a concentration of the
 CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo
 CC activated partial thromboplastin time (APTT) assay; calculating
 CC prolongation of clotting effected by the isolated protein in each of
 CC the PT and APTT assay, with respect to a baseline clotting value for
 CC each assay, where prolongation of clotting is calculated as fold
 CC elevation of clotting time relative to a baseline clotting value, where
 CC a doubling of clotting time is deemed a two-fold elevation; and
 CC calculating a PI to APTT prolongation ratio, where a ratio at least
 CC one is indicative of factor VIIa/TF inhibitory activity. The method is
 CC useful for determining if a protein has factor VIIa/TF inhibitory
 CC activity.
 CC
 CC Sequence 91 AA;
 SO
 Query Match 100.0%; Score 486; DB 20; Length 91;
 Host Local Similarity 100.0%; Pred. No. 66-39;
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KATMGCGENFYDSCGSKETKTKYKGVPEFPIHFNVPCLVWVHGVVPEGFRNK 50
 Lb 8 KATMGCGENFYDSCGSKETKTKYKGVPEFPIHFNVPCLVWVHGVVPEGFRNK 67
 QY 61 DDKCVSAEDCELDMMDFYPQTRN 84
 DB 68 ddkcvsaedceldmmdflypqrtn 91
 RESULT 5
 AAY30454
 ID AAY30454 standard; Protein; 91 AA.
 AC AAY30454;
 XX
 XX 15-NOV-1999 (first entry)
 DE Nematode extracted anticoagulant protein AcanAPc2.
 AC
 AC Nematode extracted anticoagulant protein; NAP; anticoagulant;
 KW serine protease inhibitor; NAP domain; factor VIIa/TF.
 XX
 XX *Ancylostoma caninum*.
 OS
 XX

XX	US955294 A.
XX	21-SEP-1999.
XX	
XX	19 APR-1996; 960S-0634641.
XX	
XX	17-APR-1996; 960S-0634641.
XX	18-OCT-1994; 940S-0326110.
XX	05-JUN-1995; 950S-0463465.
XX	05-JUN-1995; 950S-0465380.
XX	05-JUN-1995; 950S-0486397.
XX	05-JUN-1995; 950S-0486399.
XX	17-OCT-1995; 95W0-US13231.
XX	
XX	(GARV-) CORVAS INT INC.
XX	
XX	Bertram PM, Ganssmann YGL, Jaspers LS, Laroche YR;
XX	Larocqys MJ, Messens JHL, Moylo M, Stanssens PEH;
XX	Vlasov GP;
XX	WPI: 1999-53956/45.
XX	
XX	Screening an isolated protein for Nematode-extracted Anticoagulant
XX	protein domains
XX	
XX	Inclusionary Columns 175-176; 197pp; English.
XX	
XX	The present sequence represents a nematode extracted anticoagulant
XX	protein (NAP). The protein has activity as an anticoagulant and/or serine
XX	protease inhibitor. The protein contains at least one NAP domain which
XX	has selective inhibitory activity for factor VIIa/TF. The specification
XX	describes a method for screening an isolated protein at least one domain
XX	for factor VIIa/TF selective inhibitory activity. The method comprises
XX	determining the time to clotting effected by a concentration of the
XX	isolated protein in an ex vivo prolongation time (PT) assay and an ex vivo
XX	prolongation of clotting effected by the isolated protein in each of
XX	the PT and aPTT assay, with respect to a baseline clotting value for
XX	each assay, where prolongation of clotting is calculated as fold
XX	elevation of clotting time relative to a baseline clotting value, where
XX	a doubling of clotting time is deemed a two-fold elevation; and
XX	calculating a PT to aPTT prolongation ratio, where a ratio at least
XX	one is indicative of factor VIIa/TF inhibitory activity. The method is
XX	useful for determining if a protein has factor VIIa/TF inhibitory
XX	activity.
XX	
XX	Sequence 91 AA:
XX	
XX	Query Match: 100.0%; Score: 486; FR 20; Length 91;
XX	Host Local Similarity: 100.0%; Prod No: 6039;
XX	Matches: 84; Conservation: 0; Mutations: 0; Ticks: 0; Gaps: 0
XX	
XX	1 KATMGCHRNKYSQSGSKEDSRKNKYVEFFIETEPANVCLVWCHGVWLEDSYENK 90
XX	
XX	8 KATMQGCHRNKYSQSGSKEDSRKNKYVEFFIETEPANVCLVWCHGVWLEDSYENK 57
XX	
XX	91 DKKVSASPCFELNNKPETYPSTPN 84
XX	
XX	68 ddkvsaspcfdldmndtlyptpn 91
XX	
XX	RESULT 6
XX	AAB15346
XX	AAB15346 standard; Protein: 91 AA.
XX	
XX	AAB15346;
XX	
XX	19-DEC-2000 (first entry)
XX	
XX	A. caenium nematode-extracted anticoagulant protein AcaNAPE2.
XX	Nematode extracted and isoquantant protein; AcaNAPE2; blood clotting,

XX canine hookworm; Thrombosis; vaccine.

XX Anicylostoma caninum.

FN US6087487-A.

PN 11-JUL-2000.

PD 12-FEB-1999; 99US-0249451.

PE 17-OCT-1995; 95WO-IN13231.

PR 17-APR-1997; 97US-0809455.

PR 18-OCT-1994; 94US-0326110.

PR 05-10N-1995; 95US-0461965.

PR 05-JUN-1995; 95US-0465380.

PR 05-JUN-1995; 95US-0486397.

PR 05-JUN-1995; 95US-0486399.

XX (CORV-) CORVAS INT INC.

PA Laawereys MJ, Stanssens PEH, Jespers LS, Ganssemans YGL, Moyle M;

P1 Bergum PW, Messens JHL, Lakoche YR, Vlaask GP;

XX WP1: 2000-531359/48.

DR N-PDOB: AAV3373.

XX New cDNA molecule encoding a protein having factor Xa inhibitory

PT activity for preventing and treating blood clotting disorders -

FT comprises remacle-extracted anticoagulant protein domains -

FS disclosure; Fig 11; 197pp; English.

XX The present sequence is the Anicylostoma caninum remacle extracted

C2 anticoagulant protein AcanMP2. Proteins of this kind have been shown to

C3 bleeding. The protein can be used in blood collection tubes to aid the

CC isolation of plasma from the blood, to prevent thrombosis which may be

CC linked to the rupture of an atherosclerotic plaque, acute myocardial

CC interaction, angina, thrombotic therapy, percutaneous transluminal

CC coronary angioplasty, disseminated intravascular coagulopathy, infection,

CC cancer and septic shock, and to produce antibodies. In the latter

CC instance, the antibodies can be raised in order to detect infection by

C2 remacles. The coding sequence can also be used for this) or as

CC diagnostic tests. The protein can also be used as a vaccine against

CC remacle parasites.

XX Sequence 91 AA:

SQ

Query Match 100.0%; Score 486; DB 21; Length 91;

Best local Similarity 100.0%; Prod No 66339; C1 Indels 0; Gaps 0;

Motif:es A4: Cons-vaive; C2 Misad-mos; C3 Indels 0; Gaps 0;

VY 1 KAIMGCEENKRYEYSKSELDREKYDVADEADAEHWETIAVEWJHGVWFEHSYEAK 60

IL g kalmegceekryesysksekseldreydyadeadaeahwetiavewjhgfvwfehsyeak 67

CY 61 EIKCVSAKEPFLDNMHPYPTGN 84

DB 68 ddkcvssadeceldnmhtypatn 91

RESULT 7

ID AAV30399 standard; Protein: 82 AA.

AAV30399;

AAV30399;

DI 15-NOV-1999 (first entry)

DE Remacle extracted anticoagulant protein AcanMP442.

XX Remacle extracted anticoagulant protein, MAP, anticoagulant;

FW Remacle extracted anticoagulant protein, MAP, anticoagulant;

KM serine protease inhibitor; NAP domain; factor VIIa/TF.
 XX
 OS Ancylostoma ceylanicum.
 XX
 PN US955294-A.
 XX
 PD 21-SEP-1999.
 XX
 PF 19-APR-1996; 960S-0634641.
 XX
 PR 19-APR-1996; 960S-0634641.
 PR 18-OCT-1994; 940S-0326110.
 PR 05-JUN-1995; 950S-0461965.
 PR 05-JUN-1995; 950S-0465380.
 PR 05-JUN-1995; 950S-0463397.
 PR 05-JUN-1995; 950S-0463399.
 PR 17-OCT-1995; 95WO-0513231.
 XX
 PA (CORV-) CORVAS INT INC.
 XX
 PI Bergum PW, Gansseman YGJ, Jespers LS, Laroche YR.
 PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH.
 PI Vlasuk GP.
 XX
 DR WP1: 1999-539569/45.
 XX
 PT Screening an isolated protein for Nematode-extracted Anticoagulant
 PT Protein domains
 PS Disclousure: Column 111-112; 197pp; English.
 XX
 CC The present sequence represents a nematode extracted anticoagulant
 CC protein (NAP). The protein has activity as an anticoagulant and/or serine
 CC protease inhibitor. The protein contains at least one NAP domain which
 CC has selective inhibitory activity for factor VIIa/TF. The specification
 CC describes a method for screening an isolated protein at least one domain
 CC for factor VIIa/TF selective inhibitory activity. The method comprises
 CC determining the time to clotting effected by a concentration of the
 CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo
 CC activated partial thromboplastin time (APTT) assay; calculating
 CC prolongation of clotting effected by the isolated protein in each of
 CC the PT and APTT assay, with respect to a baseline clotting value for
 CC each assay, where prolongation of clotting is calculated as fold
 CC elevation of clotting time relative to a baseline clotting value, where
 CC a doubling of clotting time is deemed a two fold elevation; and
 CC calculating a PT to APTT prolongation ratio, where a ratio at least
 CC one is indicative of factor VIIa/TF inhibitory activity. The method is
 CC useful for determining if a protein has factor VIIa/TF inhibitory
 CC activity.
 XX
 SS Sequence 82 AA:
 SQ
 Query Match 43.6%; Score 212; DP 20; Length 82;
 Best Local Similarity 50.0%; Pred. No 3, 1e-13;
 Matches 38; Conservative 11; Mismatches 19; Indels 8; Gaps 3;
 QY 6 CGENFRVVSSTGKQVKKRYKVCPEPQVEENVAQVAVVWQDQGVLESEFYRNKRD 62
 DB 4 CGSNERGFRVNDKPYQVWQDHYKQKPKKPKKPKKPKKPKKPKKPKKPKKPKK 60
 QY 63 KCVSAEDCEIDNMDPI 78
 DB 59 scvesddceydmndf 74
 RESULT 8
 AAY30422
 ID AAY30422 standard; Protein: 82 AA.
 AC AAY10422;
 XX
 DT 15-NOV-1999 (first entry)

XX
 DE Mature nematode extracted anticoagulant protein AcenAP4d2.
 XX
 KM Nematode extracted anticoagulant protein; NAP; anticoagulant;
 KM serine protease inhibitor; NAP domain; factor VIIa/TF.
 XX
 OS Ancylostoma ceylanicum.
 XX
 PN US955294-A.
 XX
 PD 21-SEP-1999.
 XX
 PF 19-APR-1996; 960S-0634641.
 XX
 PR 19-APR-1996; 960S-0634641.
 PR 18-OCT-1994; 940S-0326110.
 PR 05-JUN-1995; 950S-0461965.
 PR 05-JUN-1995; 950S-0465380.
 PR 05-JUN-1995; 950S-0463397.
 PR 05-JUN-1995; 950S-0463399.
 PR 17-OCT-1995; 95WO-0513231.
 XX
 PA (CORV-) CORVAS INT INC.
 XX
 PI Bergum PW, Gansseman YGJ, Jespers LS, Laroche YR.
 PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH.
 PI Vlasuk GP.
 XX
 DR WP1: 1999-539569/45.
 XX
 PT Screening an isolated protein for Nematode extracted Anticoagulant
 PT Protein domains
 PS Disclousure: Columns 135-136; 197pp; English.
 XX
 CC The present sequence represents a nematode extracted anticoagulant
 CC protein (NAP). The protein has activity as an anticoagulant and/or serine
 CC protease inhibitor. The protein contains at least one NAP domain which
 CC has selective inhibitory activity for factor VIIa/TF. The specification
 CC describes a method for screening an isolated protein at least one domain
 CC for factor VIIa/TF selective inhibitory activity. The method comprises
 CC determining the time to clotting effected by a concentration of the
 CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo
 CC activated partial thromboplastin time (APTT) assay; calculating
 CC prolongation of clotting effected by the isolated protein in each of
 CC the PT and APTT assay, with respect to a baseline clotting value for
 CC each assay, where prolongation of clotting is calculated as fold
 CC elevation of clotting time relative to a baseline clotting value, where
 CC a doubling of clotting time is deemed a two fold elevation; and
 CC calculating a PT to APTT prolongation ratio, where a ratio at least
 CC one is indicative of factor VIIa/TF inhibitory activity. The method is
 CC useful for determining if a protein has factor VIIa/TF inhibitory
 CC activity.
 XX
 SS Sequence 82 AA:
 SQ
 Query Match 43.6%; Score 212; DP 20; Length 82;
 Best Local Similarity 50.0%; Pred. No 3, 1e-13;
 Matches 38; Conservative 11; Mismatches 19; Indels 8; Gaps 3;
 QY 6 CGENFRVVSSTGKQVKKRYKVCPEPQVEENVAQVAVVWQDQGVLESEFYRNKRD 62
 DB 4 CGSNERGFRVNDKPYQVWQDHYKQKPKKPKKPKKPKKPKKPKKPKKPKKPKK 60
 QY 63 KCVSAEDCEIDNMDPI 78
 DB 59 scvesddceydmndf 74
 RESULT 9
 AAB15293
 ID AAB15293 standard; Protein: 82 AA.

With receptor precursor - African clawed frog
 M:Altrorad name: very low density lipoprotein receptor, vitellogenin receptor
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 16-Jun-2000
 C:Accession: J04858
 R:Obayashi, K.; Shoji, H.; Nakamura, T.; Hashimoto, O.; Asashima, M.; Saito, H.
 Biochem. Biophys. Res. Commun. 224, 406-413, 1996
 A:Title: cDNA cloning and expression of the Xenopus laevis vitellogenin receptor.
 A:Reference number: J04858; M010:96295501
 A:Accession: J04858
 A:Molecule type: mRNA
 A:Accession: U869-00KA
 C:Comment: This receptor mediates incorporation of vitellogenin into oocytes.
 C:Superfamily: LDL receptor; EGF homology; LDL receptor ligand-binding repeat homology;
 C:Keywords: duplication, fatty acid metabolism, vitellogenin, receptor, transmembrane pr
 E:1-26/Domain: signal sequence #status predicted <MAT>
 E:27-869/Domain: WLD receptor #status predicted <MAT>
 E:27-794/Domain: extracellular #status predicted <EXT>
 E:42-66/Domain: LDL receptor ligand-binding repeat homology <LDL>
 E:71-107/Domain: LDL receptor ligand-binding repeat homology <LDL>
 E:112-134/Domain: LDL receptor ligand-binding repeat homology <LDL>
 E:153-187/Domain: LDL receptor ligand-binding repeat homology <LDL>
 E:192-228/Domain: LDL receptor ligand-binding repeat homology <LDL>
 E:248-272/Domain: LDL receptor ligand-binding repeat homology <LDL>
 E:277-311/Domain: LDL receptor ligand-binding repeat homology <LDL>
 E:317-354/Domain: LDL receptor ligand-binding repeat homology <LDL>
 E:359-393/Domain: EGF homology <EGF>
 E:399-434/Domain: EGF homology <EGF>
 E:440-479/Domain: LDL receptor WLD-receptor repeat homology <WLD>
 E:480-526/Domain: LDL receptor WLD-receptor repeat homology <WLD>
 E:526-568/Domain: LDL receptor WLD-receptor repeat homology <WLD>
 E:569-612/Domain: LDL receptor WLD-receptor repeat homology <WLD>
 E:613-655/Domain: LDL receptor WLD-receptor repeat homology <WLD>
 E:656-698/Domain: LDL receptor WLD-receptor repeat homology <WLD>
 E:707-749/Domain: EGF homology <EGF>
 E:794-815/Domain: Transmembrane #status predicted <EMB>
 E:816-869/Domain: Intracellular #status predicted <INT>
 E:840-844/Domain: coated pit mediated internalization signal
 E:840-201-777-786/Domain: EGF, partially (A60) (500-600) #status predicted
 E:839-370-386-379-381-393-399-409-405-418-429-433-707-720-716-735-737-749-751-756 bond

Query Match 19.9% Score 96.5% DB 1: Length 869;
 Best local similarity 45.4% Pred. No. 0.12;
 Matches 28; Conservative 8; Mismatches 38; Indels 5; Gaps 4;
 2 ATMGCEFNKYSKSKK-LEKKYLVVEEDPFPVWPPIVYVGHQD-CVCECECPN 59
 Db 244 ALVGFSANM-PPVSGPPIHFFWETVDAVGNKSELEINLSRQVDFGKDEGNCIH 291
 60 KKKCVSARDCLEDDMPT 78
 Db 292 GSKQTDVMDG-LDEDEL 409

RESULT 12
 142737
 gp330 protein precursor - rat
 M:Altrorad name: megalin
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
 C:Accession: T42737
 R:Saito, A.; Picotromaro, S.; Loo, A.K.C.; Farquhar, M.G.
 Proc. Natl. Acad. Sci. U.S.A. 91, 9725-9729, 1994
 A:Title: Complete cloning and sequencing of rat gp330/megalyn, a distinctive member of
 A:Reference number: A58173; M010:95024933
 A:Accession: T42737
 A:Status: preliminary; translated from cDNA/EMBL/DDDBJ
 A:Molecule type: mRNA
 A:Accession: U4660-00SA
 A:Reference number: E04-34049-NT09561853; P015561953; P0195651961
 A:Experimental source: strain Sprague-Dawley; kidney

C:Superfamily: alpha-2-macroglobulin receptor; EGF homology; LDL receptor ligand bind
 E:1-25/Domain: signal sequence #status predicted <SIG>
 E:26-460/Product: gp330 protein #status predicted <MAT>
 Query Match 19.4% Score 94.5% DB 2: Length 4660;
 Best local similarity 30.0% Pred. No. 0.75;
 Matches 30; Conservative 11; Mismatches 24; Indels 35; Gaps 8;
 6 GGNKRYDST-----GSKR---DKKRYVYR-EDDEPNV---PCLYRV... 45
 Db 256 GSNOSHHCYPRWACPSQSCISTDKVQ--DQVHCPEHDDNNVNSCPYDGMGVSV 313
 46 -----CHQ-----DCELECFYRKKDK--CVSAHDEL 72
 Db 314 TRFVWCHDLPFGSEPTFGHITNSNSPTPTDQDQ 353

RESULT 13
 507127
 chymotrypsin/elastase inhibitor - common roundworm
 C:Species: Ascaris lumbricoles (common roundworm)
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Jul-1997
 C:Accession: S07127
 R:Radin, D.R.; Penasky, R.J.; Gross, S.M.
 Arch. Biochem. Biophys. 232, 14-161, 1994
 A:Title: The inhibitors of chymotrypsin/elastase from Ascaris lumbricoles: the pr
 A:Reference number: S07127; M010:84255715
 A:Accession: S07127
 A:Status: preliminary
 A:Molecule type: protein
 A:Accession: U63-00AB
 C:Superfamily: roundworm trypsin inhibitor

Query Match 19.9% Score 92.5% DB 2: Length 63;
 Best local similarity 32.8% Pred. No. 0.93;
 Matches 22; Conservative 5; Mismatches 27; Indels 1; Gaps 4;
 5 GINFEVSTSEFTFTTGVLEHLELLELVYVFNHGLVLEGVNHLRDL 63
 Db 5 GINFEVHTCTG--CPMK--GPRP-----NIPGLMKRSPSPSGMGRKRNOK 53
 64 CVASADPT 70
 Db 54 CLPASQC 60

RESULT 14
 505572
 chymotrypsin/elastase inhibitor - common roundworm
 C:Species: Ascaris lumbricoles (common roundworm)
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
 C:Accession: S05572
 R:Radin, D.R.; Penasky, R.J.; Gross, S.M.
 Arch. Biochem. Biophys. 232, 14-161, 1994
 A:Title: The inhibitors of chymotrypsin/elastase from Ascaris lumbricoles: the pr
 A:Reference number: S07127; M010:84255715
 A:Accession: S05572
 A:Status: preliminary
 A:Molecule type: protein
 A:Accession: U63-00AB
 C:Superfamily: roundworm trypsin inhibitor

Query Match 19.0% Score 92.5% DB 2: Length 63;
 Best local similarity 29.9% Pred. No. 0.93;
 Matches 20; Conservative 7; Mismatches 27; Indels 1; Gaps 4;
 6 GGNKRYDSTSEFTFTTGVLEHLELLELVYVFNHGLVLEGVNHLRDL 63
 Db 4 GGNKRYDSTSEFTFTTGVLEHLELLELVYVFNHGLVLEGVNHLRDL 52

DE (GLYCOPROTEIN 330).
GN LRP2.
OS BATTER neurogins (Bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId:10116;
RN [1]
RF SEQUENCE FROM N.A.
RX STRAIN: SPRAIN; FAMILY: TISSUE Kidney;
RX MEDLINE: 95024033; PubMed 7317880;
RA Saito A., Petromonaco S., Luo A.K.C., Farquhar M.G.;
RI "Complete cloning and sequencing of rat gp330/megalin," a
RI distinctive member of the low density lipoprotein receptor gene
RI family";
RI Proc. Natl. Acad. Sci. U S A 91:9725-9729(1994)
RN [2]
RX PUNCTION.
RX MEDLINE: 9538696; PubMed 7544804;
RA Moestrup S.K., Gul S., Vorum H., Brogsgaard C., Hjort S.E.,
RA Morsrup K., Gliemann J., Christensen E.J.;
RI "Evidence that epithelial glycoprotein 330/megalin mediates uptake of
RI polyanionic drugs";
RI J. Clin. Invest. 96:1404-1413(1995).
RN [3]
RX TISSUE SPECIFICITY.
RX MEDLINE: 94172242; PubMed 7510321;
RA Chong G., Bachinsky D.R., Stamenkovic I., Stuckand D.R., Brown D.,
RA Andrews G., McCluskey R.T.;
RI "Organ distribution in rats of two members of the low-density
RI lipoprotein receptor gene family, gp330 and lrp2-like 2m, and the
RI receptor-associated protein (RAP).";
RI J. Histochem. Cytochem. 42:531-542(1994).
RN [4]
RX FUNCTION: BINDS PLASMINOGEN, EXTRACELLULAR MATRIX COMPONENTS,
RX PLASMINOGEN ACTIVATOR-PLASMINOGEN ACTIVATOR INHIBITOR TYPE 1
RX COMPLEX, APOLOPROTEIN E-ENRICHED BETA-VLDL, LIPOPROTEIN LIPASE,
RX LACTOPERLIN, CLUSTERIN AND CALGRIUM
RX FUNCTION: RECEPTOR-MEDIATED UPTAKE OF POLYANIONIC DRUGS SUCH AS
RX APOLIPOTEIN, AMINOGLYCOSIDES AND POLYMYXIN B.
RX SIMONIT: FORMS A MULTIMERIC COMPLEX TOGETHER WITH A RECEPTOR-
RX ASSOCIATED PROTEIN (RAP).
RX SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXPRESSED IN
RX CLATHRIN-COATED PITS. A SOLUBLE FORM IS POSSIBLY DERIVED BY
RX CLEAVAGE AT THE CELL SURFACE.
RX TISSUE SPECIFICITY: EPITHELIAL CELLS OF KIDNEY GLOMERULUS AND
RX PROXIMAL TUBULE, LUNG, EPIDIDYMIS, YOLK SAC, AMONG OTHERS.
RX SIMILARITY: CONTAINS 36 LDL-RECEPTOR CLASS A DOMAINS
RX SIMILARITY: CONTAINS 37 LDL-RECEPTOR CLASS B DOMAINS.
RX SIMILARITY: CONTAINS 3 SH2-BINDING DOMAINS.
RX SIMILARITY: CONTAINS 1 SH2-BINDING DOMAIN.
RX THIS SWISS PROTEIN ENTRY IS COPYRIGHTED. IT IS PROVIDED THROUGH A collaboration
RX between the Swiss Institute of Bioinformatics and the EMBL Outstation -
RX the European Bioinformatics Institute. There are no restrictions on its
RX use by non-profit institutions as long as its content is in no way
RX modified and this statement is not removed. Usage by and for commercial
RX entities requires a license agreement (See <http://www.isb-sib.ch/abonment/>
RX or send an email to license@sib-sib.ch)
RX EMBL: L34049; AA51369.1;
RX HSSP: P01140; IALI.
RX InterPro: IPR000033;
RX InterPro: IPR000152;
RX InterPro: IPR000561;
RX InterPro: IPR001881;
RX InterPro: IPR002172;
RX Pfam: PF000057; ldl_recept_a; 36.
RX Pfam: PF00058; ldl_recept_b; 33
RX PRINTS: PR00261; ldl_recept
RX PROSITE: PS00010; ASX_HYDROXYL; 4.
RX PROSITE: PS00022; PCP_1; 1
RX PROSITE: PS01186; PCP_2; 8.
RX PROSITE: PS01187; PCP_3A; 3.

DR PROSITE: PS01289; LDLRA_1; 31.
DR PROSITE: PS00068; LDLRA_2; 26.
KW Glycoprotein, Recept., Endocytosis, Coated pits; Transmembrane;
KW Receptor; EGF Like domain; Signal.
FT SIGNAL
FT CHAIN 1 25
FT 26 4660
FT 26 4425
FT TRANSMEM
FT 4426 4446
FT 4447 4660
FT 26 64
FT 65 105
FT 106 144
FT 145 181
FT 181 219
FT 219 258
FT 258 307
FT 307 346
FT 346 385
FT 385 435
FT 435 476
FT 476 519
FT 519 566
FT 566 611
FT 612 652
FT 652 704
FT 704 752
FT 752 793
FT 793 835
FT 835 879
FT 879 923
FT 923 1013
FT 1013 1061
FT 1061 1108
FT 1108 1146
FT 1146 1186
FT 1186 1225
FT 1225 1269
FT 1269 1308
FT 1308 1351
FT 1351 1390
FT 1390 1430
FT 1430 1479
FT 1479 1522
FT 1522 1563
FT 1563 1609
FT 1609 1654
FT 1654 1696
FT 1696 1742
FT 1742 1791
FT 1791 1832
FT 1832 1882
FT 1882 1930
FT 1930 1972
FT 1972 2013
FT 2013 2050
FT 2050 2108
FT 2108 2156
FT 2156 2201
FT 2201 2245
FT 2245 2289
FT 2289 2332
FT 2332 2384
FT 2384 2432
FT 2432 2477
FT 2477 2518
FT 2518 2562
FT 2562 2604
FT 2604 2647
FT 2647 2694
FT 2694 2739
FT 2739 2778
FT 2778 2820
FT 2820 2862
FT 2862 2903
FT 2903 2947
FT 2947 2993
FT 2993 3031
FT 3031 3072
FT 3072 3113
FT 3113 3154
FT 3154 3195
FT 3195 3236
FT 3236 3277
FT 3277 3318
FT 3318 3359
FT 3359 3400
FT 3400 3441
FT 3441 3482
FT 3482 3523
FT 3523 3564
FT 3564 3605
FT 3605 3646
FT 3646 3687
FT 3687 3728
FT 3728 3769
FT 3769 3810
FT 3810 3851
FT 3851 3892
FT 3892 3933
FT 3933 3974
FT 3974 4015
FT 4015 4056
FT 4056 4097
FT 4097 4138
FT 4138 4179
FT 4179 4220
FT 4220 4261
FT 4261 4302
FT 4302 4343
FT 4343 4384
FT 4384 4425
FT 4425 4466
FT 4466 4507
FT 4507 4548
FT 4548 4589
FT 4589 4630
FT 4630 4671
FT 4671 4712
FT 4712 4753
FT 4753 4794
FT 4794 4835
FT 4835 4876
FT 4876 4917
FT 4917 4958
FT 4958 4999
FT 4999 5040
FT 5040 5081
FT 5081 5122
FT 5122 5163
FT 5163 5204
FT 5204 5245
FT 5245 5286
FT 5286 5327
FT 5327 5368
FT 5368 5409
FT 5409 5450
FT 5450 5491
FT 5491 5532
FT 5532 5573
FT 5573 5614
FT 5614 5655
FT 5655 5696
FT 5696 5737
FT 5737 5778
FT 5778 5819
FT 5819 5860
FT 5860 5901
FT 5901 5942
FT 5942 5983
FT 5983 6024
FT 6024 6065
FT 6065 6106
FT 6106 6147
FT 6147 6188
FT 6188 6229
FT 6229 6270
FT 6270 6311
FT 6311 6352
FT 6352 6393
FT 6393 6434
FT 6434 6475
FT 6475 6516
FT 6516 6557
FT 6557 6598
FT 6598 6639
FT 6639 6680
FT 6680 6721
FT 6721 6762
FT 6762 6803
FT 6803 6844
FT 6844 6885
FT 6885 6926
FT 6926 6967
FT 6967 7008
FT 7008 7049
FT 7049 7090
FT 7090 7131
FT 7131 7172
FT 7172 7213
FT 7213 7254
FT 7254 7295
FT 7295 7336
FT 7336 7377
FT 7377 7418
FT 7418 7459
FT 7459 7500
FT 7500 7541
FT 7541 7582
FT 7582 7623
FT 7623 7664
FT 7664 7705
FT 7705 7746
FT 7746 7787
FT 7787 7828
FT 7828 7869
FT 7869 7910
FT 7910 7951
FT 7951 7992
FT 7992 8033
FT 8033 8074
FT 8074 8115
FT 8115 8156
FT 8156 8197
FT 8197 8238
FT 8238 8279
FT 8279 8320
FT 8320 8361
FT 8361 8402
FT 8402 8443
FT 8443 8484
FT 8484 8525
FT 8525 8566
FT 8566 8607
FT 8607 8648
FT 8648 8689
FT 8689 8730
FT 8730 8771
FT 8771 8812
FT 8812 8853
FT 8853 8894
FT 8894 8935
FT 8935 8976
FT 8976 9017
FT 9017 9058
FT 9058 9099
FT 9099 9140
FT 9140 9181
FT 9181 9222
FT 9222 9263
FT 9263 9304
FT 9304 9345
FT 9345 9386
FT 9386 9427
FT 9427 9468
FT 9468 9509
FT 9509 9550
FT 9550 9591
FT 9591 9632
FT 9632 9673
FT 9673 9714
FT 9714 9755
FT 9755 9796
FT 9796 9837
FT 9837 9878
FT 9878 9919
FT 9919 9960
FT 9960 10001
FT 10001 10042
FT 10042 10083
FT 10083 10124
FT 10124 10165
FT 10165 10206
FT 10206 10247
FT 10247 10288
FT 10288 10329
FT 10329 10370
FT 10370 10411
FT 10411 10452
FT 10452 10493
FT 10493 10534
FT 10534 10575
FT 10575 10616
FT 10616 10657
FT 10657 10698
FT 10698 10739
FT 10739 10780
FT 10780 10821
FT 10821 10862
FT 10862 10903
FT 10903 10944
FT 10944 10985
FT 10985 11026
FT 11026 11067
FT 11067 11108
FT 11108 11149
FT 11149 11190
FT 11190 11231
FT 11231 11272
FT 11272 11313
FT 11313 11354
FT 11354 11395
FT 11395 11436
FT 11436 11477
FT 11477 11518
FT 11518 11559
FT 11559 11600
FT 11600 11641
FT 11641 11682
FT 11682 11723
FT 11723 11764
FT 11764 11805
FT 11805 11846
FT 11846 11887
FT 11887 11928
FT 11928 11969
FT 11969 12010
FT 12010 12051
FT 12051 12092
FT 12092 12133
FT 12133 12174
FT 12174 12215
FT 12215 12256
FT 12256 12297
FT 12297 12338
FT 12338 12379
FT 12379 12420
FT 12420 12461
FT 12461 12502
FT 12502 12543
FT 12543 12584
FT 12584 12625
FT 12625 12666
FT 12666 12707
FT 12707 12748
FT 12748 12789
FT 12789 12830
FT 12830 12871
FT 12871 12912
FT 12912 12953
FT 12953 12994
FT 12994 13035
FT 13035 13076
FT 13076 13117
FT 13117 13158
FT 13158 13199
FT 13199 13240
FT 13240 13281
FT 13281 13322
FT 13322 13363
FT 13363 13404
FT 13404 13445
FT 13445 13486
FT 13486 13527
FT 13527 13568
FT 13568 13609
FT 13609 13650
FT 13650 13691
FT 13691 13732
FT 13732 13773
FT 13773 13814
FT 13814 13855
FT 13855 13896
FT 13896 13937
FT 13937 13978
FT 13978 14019
FT 14019 14060
FT 14060 14101
FT 14101 14142
FT 14142 14183
FT 14183 14224
FT 14224 14265
FT 14265 14306
FT 14306 14347
FT 14347 14388
FT 14388 14429
FT 14429 14470
FT 14470 14511
FT 14511 14552
FT 14552 14593
FT 14593 14634
FT 14634 14675
FT 14675 14716
FT 14716 14757
FT 14757 14798
FT 14798 14839
FT 14839 14880
FT 14880 14921
FT 14921 14962
FT 14962 15003
FT 15003 15044
FT 15044 15085
FT 15085 15126
FT 15126 15167
FT 15167 15208
FT 15208 15249
FT 15249 15290
FT 15290 15331
FT 15331 15372
FT 15372 15413
FT 15413 15454
FT 15454 15495
FT 15495 15536
FT 15536 15577
FT 15577 15618
FT 15618 15659
FT 15659 15700
FT 15700 15741
FT 15741 15782
FT 15782 15823
FT 15823 15864
FT 15864 15905
FT 15905 15946
FT 15946 15987
FT 15987 16028
FT 16028 16069
FT 16069 16110
FT 16110 16151
FT 16151 16192
FT 16192 16233
FT 16233 16274
FT 16274 16315
FT 16315 16356
FT 16356 16397
FT 16397 16438
FT 16438 16479
FT 16479 16520
FT 16520 16561
FT 16561 16602
FT 16602 16643
FT 16643 16684
FT 16684 16725
FT 16725 16766
FT 16766 16807
FT 16807 16848
FT 16848 16889
FT 16889 16930
FT 16930 16971
FT 16971 17012
FT 17012 17053
FT 17053 17094
FT 17094 17135
FT 17135 17176
FT 17176 17217
FT 17217 17258
FT 17258 17299
FT 17299 17340
FT 17340 17381
FT 17381 17422
FT 17422 17463
FT 17463 17504
FT 17504 17545
FT 17545 17586
FT 17586 17627
FT 17627 17668
FT 17668 17709
FT 17709 17750
FT 17750 17791
FT 17791 17832
FT 17832 17873
FT 17873 17914
FT 17914 17955
FT 17955 17996
FT 17996 18037
FT 18037 18078
FT 18078 18119
FT 18119 18160
FT 18160 18201
FT 18201 18242
FT 18242 18283
FT 18283 18324
FT 18324 18365
FT 18365 18406
FT 18406 18447
FT 18447 18488
FT 18488 18529
FT 18529 18570
FT 18570 18611
FT 18611 18652
FT 18652 18693
FT 18693 18734
FT 18734 18775
FT 18775 18816
FT 18816 18857
FT 18857 18898
FT 18898 18939
FT 18939 18980
FT 18980 19021
FT 19021 19062
FT 19062 19103
FT 19103 19144
FT 19144 19185
FT 19185 19226
FT 19226 19267
FT 19267 19308
FT 19308 19349
FT 19349 19390
FT 19390 19431
FT 19431 19472
FT 19472 19513
FT 19513 19554
FT 19554 19595
FT 19595 19636
FT 19636 19677
FT 19677 19718
FT 19718 19759
FT 19759 19800
FT 19800 19841
FT 19841 19882
FT 19882 19923
FT 19923 19964
FT 19964 20005
FT 20005 20046
FT 20046 20087
FT 20087 20128
FT 20128 20169
FT 20169 20210
FT 20210 20251
FT 20251 20292
FT 20292 20333
FT 20333 20374
FT 20374 20415
FT 20415 20456
FT 20456 20497
FT 20497 20538
FT 20538 20579
FT 20579 20620
FT 20620 20661
FT 20661 20702
FT 20702 20743
FT 20743 20784
FT 20784 20825
FT 20825 20866
FT 20866 20907
FT 20907 20948
FT 20948 20989
FT 20989 21030
FT 21030 21071
FT 21071 21112
FT 21112 21153
FT 21153 21194
FT 21194 21235
FT 21235 21276
FT 21276 21317
FT 21317 21358
FT 21358 21399
FT 21399 21440
FT 21440 21481
FT 21481 21522
FT 21522 21563
FT 21563 21604
FT 21604 21645
FT 21645 21686
FT 21686 21727
FT 21727 21768
FT 21768 21809
FT 21809 21850
FT 21850 21891
FT 21891 21932
FT 21932 21973
FT 21973 22014
FT 22014 22055
FT 22055 22096
FT 22096 22137
FT 22137 22178
FT 22178 22219
FT 22219 22260
FT 22260 22301
FT 22301 22342
FT 22342 22383
FT 22383 22424
FT 22424 22465
FT 22465 22506
FT 22506 22547
FT 22547 22588
FT 22588 22629
FT 22629 22670
FT 22670 22711
FT 22711 22752
FT 22752 22793
FT 22793 22834
FT 22834 22875
FT 22875 22916
FT 22916 22957
FT 22957 22998
FT 22998 23039
FT 23039 23080
FT 23080 23121
FT 23121 23162
FT 23162 23203
FT 23203 23244
FT 23244 23285
FT 23285 23326
FT 23326 23367
FT 23367 23408
FT 23408 23449
FT 23449 23490
FT 23490 23531
FT 23531 23572
FT 23572 23613
FT 23613 23654
FT 23654 23695
FT 23695 23736
FT 23736 23777
FT 23777 23818
FT 23818 23859
FT 23859 23900
FT 23900 23941
FT 23941 23982
FT 23982 24023
FT 24023 24064
FT 24064 24105
FT 24105 24146
FT 24146 24187
FT 24187 24228
FT 24228 24269
FT 24269 24310
FT 24310 24351
FT 24351 24392
FT 24392 24433
FT 24433 24474
FT 24474 24515
FT 24515 24556
FT 24556 24597
FT 24597 24638
FT 24638 24679
FT 24679 24720
FT 24720 24761
FT 24761 24802
FT 24802 24843
FT 24843 24884
FT 24884 24925
FT 24925 24966
FT 24966 25007
FT 25007 25048
FT 25048 25089
FT 25089 25130
FT 25130 25171
FT 25171 25212
FT 25212 25253
FT 25253 25294
FT 25294 25335
FT 25335 25376
FT 25376 25417
FT 25417 25458
FT 25458 25499
FT 25499 25540
FT 25540 25581
FT 25581 25622
FT 25622 25663
FT 25663 25704
FT 25704 25745
FT 25745 25786
FT 25786 25827
FT 25827 25868
FT 25868 25909
FT 25909 25950
FT 25950 25991
FT 25991 26032
FT 26032 26073
FT 26073 26114
FT 26114 26155
FT 26155 26196
FT 26196 26237
FT 26237 26278
FT 26278 26319
FT 26319 26360
FT 26360 26401
FT 26401 26442
FT 26442 26483
FT 26483 26524
FT 26524 26565
FT 26565 26606
FT 26606 26647
FT 26647 26688
FT 26688 26729
FT 26729 26770
FT 26770 26811
FT 26811 26852
FT 26852 26893
FT 26893 26934
FT 26934 26975
FT 26975 27016
FT 27016 27057
FT 27057 27098
FT 27098 27139
FT 27139 27180
FT 27180 27221
FT 27221 27262
FT 27262 27303
FT 27303 27344
FT 27344 27385
FT 27385 27426
FT 27426 27467
FT 27467 27508
FT 27508 27549
FT 27549 27590
FT 27590 27631
FT 27631 27672
FT 27672 27713
FT 27713 27754
FT 27754 27795
FT 27795 27836
FT 27836 27877
FT 27877 27918
FT 27918 27959
FT 27959 28000
FT 28000 28041
FT 28041 28082
FT 28082 28123
FT 28123 28164
FT 28164 28205
FT 28205 28246
FT 28246 28287
FT 28287 28328
FT 28328 28369
FT 28369 28410
FT 28410 28451
FT 28451 28492
FT 28492 28533
FT 28533 28574
FT 28574 28615
FT 28615 28656
FT 28656

CC -1- SIMILARITY: CONTAINS 35 LDL-RECEPTOR CLASS A DOMAINS.
 CC -1- SIMILARITY: CONTAINS 17 EGF-LIKE DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ist.slb.ch/announcements>
 CC or send an email to license@ist.slb.ch).
 CC
 DR EMBL: M96150; AAA28105.1; -
 DR EMBL: Z73907; CAA98124.1; -
 DR HSSP: P07204; 1EGF.
 DR Wormpep: P29011.1; CE05765.
 DR InterPro: IPR000933; -
 DR InterPro: IPR000561; -
 DR InterPro: IPR001881; -
 DR InterPro: IPR002172; -
 DR Pfam: PF00057; ldl_recept_a; 35
 DR Pfam: PF00058; ldl_recept_b; 26.
 DR PRINTS: PR00261; LDLRECEPTOR.
 DR PROSITE: PS00010; ASX_HYDROXYL; 6.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; 3.
 DR PROSITE: PS01187; EGF_CA; 3.
 DR PROSITE: PS01209; LDLRA_1; 27.
 DR PROSITE: PS00068; LDLRA_2; 34.
 DR Receptor: Transmembrane; Repeat: Endocytosis; Glycoprotein;
 KM Signal: Calcium binding; EGF like domain; Coated pits
 FT SIGNAL 1 18
 FT CHAIN 19 4753
 FT
 FT TRANSMEM 19 4570
 FT 4571 4596
 FT 4597 4753
 FT DOMAIN 51 89
 FT 90 133
 FT 136 177
 FT 180 220
 FT 221 259
 FT 260 298
 FT 299 337
 FT 338 368
 FT 369 386
 FT 387 404
 FT 405 422
 FT 423 440
 FT 441 458
 FT 459 476
 FT 477 494
 FT 495 512
 FT 513 530
 FT 531 548
 FT 549 566
 FT 567 584
 FT 585 602
 FT 603 620
 FT 621 638
 FT 639 656
 FT 657 674
 FT 675 692
 FT 693 710
 FT 711 728
 FT 729 746
 FT 747 764
 FT 765 782
 FT 783 800
 FT 801 818
 FT 819 836
 FT 837 854
 FT 855 872
 FT 873 890
 FT 891 908
 FT 909 926
 FT 927 944
 FT 945 962
 FT 963 980
 FT 981 998
 FT 999 1016
 FT 1017 1034
 FT 1035 1052
 FT 1053 1070
 FT 1071 1088
 FT 1089 1106
 FT 1107 1124
 FT 1125 1142
 FT 1143 1160
 FT 1161 1178
 FT 1179 1196
 FT 1197 1214
 FT 1215 1232
 FT 1233 1250
 FT 1251 1268
 FT 1269 1286
 FT 1287 1304
 FT 1305 1322
 FT 1323 1340
 FT 1341 1358
 FT 1359 1376
 FT 1377 1394
 FT 1395 1412
 FT 1413 1430
 FT 1431 1448
 FT 1449 1466
 FT 1467 1484
 FT 1485 1502
 FT 1503 1520
 FT 1521 1538
 FT 1539 1556
 FT 1557 1574
 FT 1575 1592
 FT 1593 1610
 FT 1611 1628
 FT 1629 1646
 FT 1647 1664
 FT 1665 1682
 FT 1683 1700
 FT 1701 1718
 FT 1719 1736
 FT 1737 1754
 FT 1755 1772
 FT 1773 1790
 FT 1791 1808
 FT 1809 1826
 FT 1827 1844
 FT 1845 1862
 FT 1863 1880
 FT 1881 1898
 FT 1899 1916
 FT 1917 1934
 FT 1935 1952
 FT 1953 1970
 FT 1971 1988
 FT 1989 2006
 FT 2007 2024
 FT 2025 2042
 FT 2043 2060
 FT 2061 2078
 FT 2079 2096
 FT 2097 2114
 FT 2115 2132
 FT 2133 2150
 FT 2151 2168
 FT 2169 2186
 FT 2187 2204
 FT 2205 2222
 FT 2223 2240
 FT 2241 2258
 FT 2259 2276
 FT 2277 2294
 FT 2295 2312
 FT 2313 2330
 FT 2331 2348
 FT 2349 2366
 FT 2367 2384
 FT 2385 2402
 FT 2403 2420
 FT 2421 2438
 FT 2439 2456
 FT 2457 2474
 FT 2475 2492
 FT 2493 2510
 FT 2511 2528
 FT 2529 2546
 FT 2547 2564
 FT 2565 2582
 FT 2583 2600
 FT 2601 2618
 FT 2619 2636
 FT 2637 2654
 FT 2655 2672
 FT 2673 2690
 FT 2691 2708
 FT 2709 2726
 FT 2727 2744
 FT 2745 2762
 FT 2763 2780
 FT 2781 2798
 FT 2799 2816
 FT 2817 2834
 FT 2835 2852
 FT 2853 2870
 FT 2871 2888
 FT 2889 2906
 FT 2907 2924
 FT 2925 2942
 FT 2943 2960
 FT 2961 2978
 FT 2979 2996
 FT 2997 3014
 FT 3015 3032
 FT 3033 3050
 FT 3051 3068
 FT 3069 3086
 FT 3087 3104
 FT 3105 3122
 FT 3123 3140
 FT 3141 3158
 FT 3159 3176
 FT 3177 3194
 FT 3195 3212
 FT 3213 3230
 FT 3231 3248
 FT 3249 3266
 FT 3267 3284
 FT 3285 3302
 FT 3303 3320
 FT 3321 3338
 FT 3339 3356
 FT 3357 3374
 FT 3375 3392
 FT 3393 3410
 FT 3411 3428
 FT 3429 3446
 FT 3447 3464
 FT 3465 3482
 FT 3483 3500
 FT 3501 3518
 FT 3519 3536
 FT 3537 3554
 FT 3555 3572
 FT 3573 3590
 FT 3591 3608
 FT 3609 3626
 FT 3627 3644
 FT 3645 3662
 FT 3663 3680
 FT 3681 3698
 FT 3699 3716
 FT 3717 3734
 FT 3735 3752
 FT 3753 3770
 FT 3771 3788
 FT 3789 3806
 FT 3807 3824
 FT 3825 3842
 FT 3843 3860
 FT 3861 3878
 FT 3879 3896
 FT 3897 3914
 FT 3915 3932
 FT 3933 3950
 FT 3951 3968
 FT 3969 3986
 FT 3987 4004
 FT 4005 4022
 FT 4023 4040
 FT 4041 4058
 FT 4059 4076
 FT 4077 4094
 FT 4095 4112
 FT 4113 4130
 FT 4131 4148
 FT 4149 4166
 FT 4167 4184
 FT 4185 4202
 FT 4203 4220
 FT 4221 4238
 FT 4239 4256
 FT 4257 4274
 FT 4275 4292
 FT 4293 4310
 FT 4311 4328
 FT 4329 4346
 FT 4347 4364
 FT 4365 4382
 FT 4383 4400
 FT 4401 4418
 FT 4419 4436
 FT 4437 4454
 FT 4455 4472
 FT 4473 4490
 FT 4491 4508
 FT 4509 4526
 FT 4527 4544
 FT 4545 4562
 FT 4563 4580
 FT 4581 4598
 FT 4599 4616
 FT 4617 4634
 FT 4635 4652
 FT 4653 4670
 FT 4671 4688
 FT 4689 4706
 FT 4707 4724
 FT 4725 4742
 FT 4743 4760
 FT 4761 4778
 FT 4779 4796
 FT 4797 4814
 FT 4815 4832
 FT 4833 4850
 FT 4851 4868
 FT 4869 4886
 FT 4887 4904
 FT 4905 4922
 FT 4923 4940
 FT 4941 4958
 FT 4959 4976
 FT 4977 4994
 FT 4995 5012
 FT 5013 5030
 FT 5031 5048
 FT 5049 5066
 FT 5067 5084
 FT 5085 5102
 FT 5103 5120
 FT 5121 5138
 FT 5139 5156
 FT 5157 5174
 FT 5175 5192
 FT 5193 5210
 FT 5211 5228
 FT 5229 5246
 FT 5247 5264
 FT 5265 5282
 FT 5283 5300
 FT 5301 5318
 FT 5319 5336
 FT 5337 5354
 FT 5355 5372
 FT 5373 5390
 FT 5391 5408
 FT 5409 5426
 FT 5427 5444
 FT 5445 5462
 FT 5463 5480
 FT 5481 5498
 FT 5499 5516
 FT 5517 5534
 FT 5535 5552
 FT 5553 5570
 FT 5571 5588
 FT 5589 5606
 FT 5607 5624
 FT 5625 5642
 FT 5643 5660
 FT 5661 5678
 FT 5679 5696
 FT 5697 5714
 FT 5715 5732
 FT 5733 5750
 FT 5751 5768
 FT 5769 5786
 FT 5787 5804
 FT 5805 5822
 FT 5823 5840
 FT 5841 5858
 FT 5859 5876
 FT 5877 5894
 FT 5895 5912
 FT 5913 5930
 FT 5931 5948
 FT 5949 5966
 FT 5967 5984
 FT 5985 6002
 FT 6003 6020
 FT 6021 6038
 FT 6039 6056
 FT 6057 6074
 FT 6075 6092
 FT 6093 6110
 FT 6111 6128
 FT 6129 6146
 FT 6147 6164
 FT 6165 6182
 FT 6183 6200
 FT 6201 6218
 FT 6219 6236
 FT 6237 6254
 FT 6255 6272
 FT 6273 6290
 FT 6291 6308
 FT 6309 6326
 FT 6327 6344
 FT 6345 6362
 FT 6363 6380
 FT 6381 6398
 FT 6399 6416
 FT 6417 6434
 FT 6435 6452
 FT 6453 6470
 FT 6471 6488
 FT 6489 6506
 FT 6507 6524
 FT 6525 6542
 FT 6543 6560
 FT 6561 6578
 FT 6579 6596
 FT 6597 6614
 FT 6615 6632
 FT 6633 6650
 FT 6651 6668
 FT 6669 6686
 FT 6687 6704
 FT 6705 6722
 FT 6723 6740
 FT 6741 6758
 FT 6759 6776
 FT 6777 6794
 FT 6795 6812
 FT 6813 6830
 FT 6831 6848
 FT 6849 6866
 FT 6867 6884
 FT 6885 6902
 FT 6903 6920
 FT 6921 6938
 FT 6939 6956
 FT 6957 6974
 FT 6975 6992
 FT 6993 7010
 FT 7011 7028
 FT 7029 7046
 FT 7047 7064
 FT 7065 7082
 FT 7083 7100
 FT 7101 7118
 FT 7119 7136
 FT 7137 7154
 FT 7155 7172
 FT 7173 7190
 FT 7191 7208
 FT 7209 7226
 FT 7227 7244
 FT 7245 7262
 FT 7263 7280
 FT 7281 7298
 FT 7299 7316
 FT 7317 7334
 FT 7335 7352
 FT 7353 7370
 FT 7371 7388
 FT 7389 7406
 FT 7407 7424
 FT 7425 7442
 FT 7443 7460
 FT 7461 7478
 FT 7479 7496
 FT 7497 7514
 FT 7515 7532
 FT 7533 7550
 FT 7551 7568
 FT 7569 7586
 FT 7587 7604
 FT 7605 7622
 FT 7623 7640
 FT 7641 7658
 FT 7659 7676
 FT 7677 7694
 FT 7695 7712
 FT 7713 7730
 FT 7731 7748
 FT 7749 7766
 FT 7767 7784
 FT 7785 7802
 FT 7803 7820
 FT 7821 7838
 FT 7839 7856
 FT 7857 7874
 FT 7875 7892
 FT 7893 7910
 FT 7911 7928
 FT 7929 7946
 FT 7947 7964
 FT 7965 7982
 FT 7983 8000
 FT 8001 8018
 FT 8019 8036
 FT 8037 8054
 FT 8055 8072
 FT 8073 8090
 FT 8091 8108
 FT 8109 8126
 FT 8127 8144
 FT 8145 8162
 FT 8163 8180
 FT 8181 8198
 FT 8199 8216
 FT 8217 8234
 FT 8235 8252
 FT 8253 8270
 FT 8271 8288
 FT 8289 8306
 FT 8307 8324
 FT 8325 8342
 FT 8343 8360
 FT 8361 8378
 FT 8379 8396
 FT 8397 8414
 FT 8415 8432
 FT 8433 8450
 FT 8451 8468
 FT 8469 8486
 FT 8487 8504
 FT 8505 8522
 FT 8523 8540
 FT 8541 8558
 FT 8559 8576
 FT 8577 8594
 FT 8595 8612
 FT 8613 8630
 FT 8631 8648
 FT 8649 8666
 FT 8667 8684
 FT 8685 8702
 FT 8703 8720
 FT 8721 8738
 FT 8739 8756
 FT 8757 8774
 FT 8775 8792
 FT 8793 8810
 FT 8811 8828
 FT 8829 8846
 FT 8847 8864
 FT 8865 8882
 FT 8883 8900
 FT 8901 8918
 FT 8919 8936
 FT 8937 8954
 FT 8955 8972
 FT 8973 8990
 FT 8991 9008
 FT 9009 9026
 FT 9027 9044
 FT 9045 9062
 FT 9063 9080
 FT 9081 9098
 FT 9099 9116
 FT 9117 9134
 FT 9135 9152
 FT 9153 9170
 FT 9171 9188
 FT 9189 9206
 FT 9207 9224
 FT 9225 9242
 FT 9243 9260
 FT 9261 9278
 FT 9279 9296
 FT 9297 9314
 FT 9315 9332
 FT 9333 9350
 FT 9351 9368
 FT 9369 9386
 FT 9387 9404
 FT 9405 9422
 FT 9423 9440
 FT 9441 9458
 FT 9459 9476
 FT 9477 9494
 FT 9495 9512
 FT 9513 9530
 FT 9531 9548
 FT 9549 9566
 FT 9567 9584
 FT 9585 9602
 FT 9603 9620
 FT 9621 9638
 FT 9639 9656
 FT 9657 9674
 FT 9675 9692
 FT 9693 9710
 FT 9711 9728
 FT 9729 9746
 FT 9747 9764
 FT 9765 9782
 FT 9783 9800
 FT 9801 9818
 FT 9819 9836
 FT 9837 9854
 FT 9855 9872
 FT 9873 9890
 FT 9891 9908
 FT 9909 9926
 FT 9927 9944
 FT 9945 9962
 FT 9963 9980
 FT 9981 9998
 FT 9999 1016
 FT 1017 1034
 FT 1035 1052
 FT 1053 1070
 FT 1071 1088
 FT 1089 1106
 FT 1107 1124
 FT 1125 1142
 FT 1143 1160
 FT 1161 1178
 FT 1179 1196
 FT 1197 1214
 FT 1215 1232
 FT 1233 1250
 FT 1251 1268
 FT 1269 1286
 FT 1287 1304
 FT 1305 1322
 FT 1323 1340
 FT 1341 1358
 FT 1359 1376
 FT 1377 1394
 FT 1395 1412
 FT 1413 1430
 FT 1431 1448
 FT 1449 1466
 FT 1467 1484
 FT 1485 1502
 FT 1503 1520
 FT 1521 1538
 FT 1539 1556
 FT 1557 1574
 FT 1575 1592
 FT 1593 1610
 FT 1611 1628
 FT 1629 1646
 FT 1647 1664
 FT 1665 1682
 FT 1683 1700
 FT 1701 1718
 FT 1719 1736
 FT 1737 1754
 FT 1755 1772
 FT 1773 1790
 FT 1791 1808
 FT 1809 1826
 FT 1827 1844
 FT 1845 1862
 FT 1863 1880
 FT 1881 1898
 FT 1899 1916
 FT 1917 1934
 FT 1935 1952
 FT 1953 1970
 FT 1971 1988
 FT 1989 2006
 FT 2007 2024
 FT 2025 2042
 FT 2043 2060
 FT 2061 2078
 FT 2079 2096
 FT 2097 2114
 FT 2115 2132
 FT 2133 2150
 FT 2151 2168
 FT 2169 2186
 FT 2187 2204
 FT 2205 2222
 FT 2223 2240
 FT 2241 2258
 FT 2259 2276
 FT 2277 2294
 FT 2295 2312
 FT 2313 2330
 FT 2331 2348
 FT 2349 2366
 FT 2367 2384
 FT 2385 2402
 FT 2403 2420
 FT 2421 2438
 FT 2439 2456
 FT 2457 2474
 FT 2475 2492
 FT 2493 2510
 FT 2511 2528
 FT 2529 2546
 FT 2547 2564
 FT 2565 2582
 FT 2583 2600
 FT 2601 2618
 FT 2619 2636
 FT 2637 2654
 FT 2655 2672
 FT 2673 2690
 FT 2691 2708
 FT 2709 2726
 FT 2727 2744
 FT 2745 2762
 FT 2763 2780
 FT 2781 2798
 FT 2799 2816
 FT 2817 2834
 FT 2835 2852
 FT 2853 2870
 FT 2871 2888
 FT 2889 2906
 FT 2907 2924
 FT 2925 2942
 FT 2943 2960
 FT 2961 2978
 FT 2979 2996
 FT 2997 3014
 FT 3015 3032
 FT 3033 3050
 FT 3051 3068
 FT 3069 3086
 FT 3087 3104
 FT 3105 3122
 FT 3123 3140
 FT 3141 3158
 FT 3159 3176
 FT 3177 3194
 FT 3195 3212
 FT 3213 3230
 FT 3231 3248
 FT 3249 3266
 FT 3267 3284
 FT 3285 3302
 FT 3303 3320
 FT 3321 3338
 FT 3339 3356
 FT 3357 3374
 FT 3375 3392
 FT 3393 3410
 FT 3411 3428
 FT 3429 3446
 FT 3447 3464
 FT 3465 3482
 FT 3483 3500
 FT 3501 3518
 FT 3519 3536
 FT 3537 3554
 FT 3555 3572
 FT 3573 3590
 FT 3591 3608
 FT 3609 3626
 FT 3627 3644
 FT 3645 3662
 FT 3663 3680
 FT 3681 3698
 FT 3699 3716
 FT 3717 3734
 FT 3735 3752
 FT 3753 3770
 FT 3771 3788
 FT 3789 3806
 FT 3807 3824
 FT 3825 3842
 FT 3843 3860
 FT 3861 3878
 FT 3879 3896
 FT 3897 3914
 FT 3915 3932
 FT 3933 3950
 FT 3951 3968
 FT 3969 3986
 FT 3987 4004
 FT 4005 4022
 FT 4023 4040
 FT 4041 4058
 FT 4059 4076
 FT 4077 4094
 FT 4095 4112
 FT 4113 4130
 FT 4131 4148
 FT 4149 4166
 FT 4167 4184
 FT 4185 4202
 FT 4203 4220
 FT 4221 4238
 FT 4239 4256
 FT 4257 4274
 FT 4275 4292
 FT 4293 4310
 FT 4311 4328
 FT 4329 4346
 FT 4347 4364
 FT 4365 4382
 FT 4383 4400
 FT 4401 4418
 FT 4419 4436
 FT 4437 4454
 FT 4455 4472
 FT 4473 4490
 FT 4491 4508
 FT 4509 4526
 FT 4527 4544
 FT 4545 4562
 FT 4563 4580
 FT 4581 4598
 FT 4599 4616
 FT 4617 4634
 FT 4635 4652
 FT 4653 4670
 FT 4671 4688
 FT 4689 4706
 FT 4707 4724
 FT 4725 4742
 FT 4743 4760
 FT 4761 4778
 FT 4779 4796
 FT 4797 4814
 FT 4815 4832
 FT 4833 4850
 FT 4851 4868
 FT 4869 4886
 FT 4887 4904
 FT 4905 4922
 FT 4923 4940
 FT 4941 4958
 FT 4959 4976
 FT 4977 4994
 FT 4995 5012
 FT 5013 5030
 FT 5031 5048
 FT 5049 5066
 FT 5067 5084
 FT 5085 5102
 FT 5103 5120


```

FT DISULFID 1408 1423 RV SIMILARITY
FT DISULFID 1423 1435 RV SIMILARITY
FT DISULFID 1441 1451 BV SIMILARITY
FT DISULFID 1447 1460 BV SIMILARITY
FT DISULFID 1462 1475 BV SIMILARITY
FT DISULFID 1751 1760 BV SIMILARITY
FT DISULFID 1756 1770 BV SIMILARITY
FT DISULFID 1772 1785 BV SIMILARITY
FT DISULFID 2084 2095 BV SIMILARITY
FT DISULFID 2091 2105 BV SIMILARITY
FT DISULFID 2107 2119 BV SIMILARITY
FT DISULFID 2400 2415 BV SIMILARITY
FT DISULFID 2411 2426 BV SIMILARITY
FT DISULFID 2428 2438 BV SIMILARITY
FT DISULFID 2732 2743 BV SIMILARITY
FT DISULFID 2739 2750 BV SIMILARITY
FT DISULFID 2761 2779 BV SIMILARITY
FT DISULFID 2792 2805 BV SIMILARITY
FT DISULFID 2803 2818 BV SIMILARITY
FT DISULFID 2812 2829 BV SIMILARITY
FT DISULFID 2834 2846 BV SIMILARITY
FT DISULFID 2841 2859 BV SIMILARITY
FT DISULFID 2859 2866 BV SIMILARITY
FT EYEPTIT 2874 2896 BV SIMILARITY
FT DISULFID 2881 2899 BV SIMILARITY
FT DISULFID 2893 2912 BV SIMILARITY
FT DISULFID 2919 2932 BV SIMILARITY

Query Match 19.0% Score 92.5 DR 1: Length 4753.
Best Local Similarity 25.0% Prod No 631. Labels 41: Caps 5
Matches 26: Conservative 11, Mismatches 32, Gaps 41

OY 7 GENETICSCTFEC----- LPTVYCYCEYCEYCEYCEY 41
D6 1174 GSPKRIEMGNNAIPAAWGSANRSPILYMLGKLNINR-YKQIDQPEEPYQYSAIG 1242
OY 42 -----LYVNGQNVCEPPEYRREKREKQ-VSAHECHLNMAPIYNG 91
D6 1233 EFYANTQCYVQVQNVHNGVQSTGKELR---SPHNSAVIAESVPIYKQVSSG 1201

RESULT 5
WVF_CANFA STANDARD: FFT: 2813 AA.
AC Q292661.028111.000814)
DT 01-NOV-1997 (Rel 35, created)
BT 01-NOV-2008 (Rel 40, last sequence update)
DI 01-OCT-2000 (Rel 40, last annotation update)
DE VON WILDBRAND FACTOR PRECURSOR (WVF).
GN FBWVF OR WVF.
OS Caris familiaris (dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi,
OC Mammalia; Eutheria; Carnivora; fissipedia; Canidae; Canis.
OX NCBI_Taxid=9615;
KN 11)
KN SEQUENCE FROM N.A.
RP STROY S I, Shibuya H, Noneman P I, Holzhauser J, Mohammad T H,
RA Johnson G S;
RA Submitted (Aug-1996) to the EMBL/GenBank/DBPI databases.
RN 131
RP SEQUENCE FROM N.A.
PA Venter P I, Li T, Vythasingam-Nathan V, Brewer C I, Stahl W P;
RI "Complete sequence of the structural gene for canine von Willebrand
RI factor and identification of a mutation causing Scottish Terrier von
RI Willebrand's disease".
RN 14)
RN Submitted (Oct-1998) to the EMBL/GenBank/DBPI databases.
RP SEQUENCE OF 1234-1669 FROM N.A
NC TISSUE=Blood;
```

PA	Maintenon T.	Christopheerson P.A., Krueter P.A., Muller-pomery P.F.,
R1	"The cause you willbrand factor gene; sequence and expression of	
R1	a region encoding the glycoprotein ID/IX binding domain.";	
RL	Submitted (JAN-1984) to the EMBL/Genbank/Jolly databases	
C1	-1 FUNCTION: IMPORTANT IN THE MAINTENANCE OF HEMOSTASIS, IT	
C1	PARTICIPATES IN PLATELET-VESSEL WALL INTERACTIONS BY FORMING A	
C1	NONVALENT COMPLEX WITH COAGULATION FACTOR VIII AT THE SITE OF	
C1	VASCULAR INJURY. (BY SIMILARITY).	
C1	-1 SUBUNIT: MULTIMERIC (BY SIMILARITY).	
C1	-1 TISSUE SPECIFICITY: BLOOD.	
C1	-1 PWM: ALL CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN OR	
C1	INTERCHAIN DISULFIDE BONDS (BY SIMILARITY).	
C1	-1 SIMILARITY: CONTAINS 3 VWFA DOMAINS.	
C1	-1 SIMILARITY: CONTAINS 3 VWFC DOMAINS.	
C1	-1 SIMILARITY: CONTAINS 4 VWFD DOMAINS.	
C1	-1 SIMILARITY: CONTAINS A C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).	
C1	-1 SIMILARITY: SOME TO SICKLE-CELL HEMOGLOBIN.	
C1	-----	
C1	This SWISS-Prot entry is reprinted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL databank on the European Bioinformatics Institute. There are no restrictions on its use by non profit institutions as long as its content is in no way modified and this statement is not removed (usage by and for commercial motifs requires a license agreement (http://www.isb-sil.ch/motifs/) or send an email to license@isb-sib.ch)	
CC	EMBL: L76227, AAC05149.1; -	
DR	EMBL: L16903; AAA30903.1; -	
DR	EMBL: AF094154, AAD04419.1; -	
DR	EMBL: U06445; AA094/b6.1; -	
DR	HSCST, F04275; IATZ.	
DR	InterPro: IPR000359; -	
IP	In-Citro: I40001007; -	
DR	InterPro: IPR001846; -	
IP	In-Citro: IPR002035; -	
DR	InterPro: IPR002919; -	
DR	FtAm: PF000907; Gys_knot; 1.	
DR	FtAm: PF01826; Ttl; 4.	
IP	FtAm: PF00042; Vwa; 3.	
DR	FtAm: PF00093; Vwc; 3.	
DR	FtAm: PF00044; vwd; 4	
DR	PRINTS: PR00453; VWPADOMAIN	
DR	PROSITE: PS01185; CTCK_1; 1	
DR	PROSITE: PS01223; CTCK_2; 1.	
IP	PROSITE: PS00334; VWFA; 3.	
DR	PROSITE: PS01208; VWFC; 3.	
DR	Related proteins: Pfam: Pfam-104; -	
KW	Plasma: Endothelial Cell; Fibro; Cell adhesion; Signal.	
FT	SIGNAL 1 22 Glycophorin, Extracellular matrix;	
FT	PROPEP 23 763 BY SIMILARITY.	
FT	CHAIN 664 2813 VON WILLEBRAND FACTOR.	
FT	DOMAIN 35 179 VFWD 1.	
FT	DOMAIN 488 541 VFWD 2.	
FT	DOMAIN 764 787 AMINO-TERMINAL.	
FT	DOMAIN 788 833 EL.	
FT	DOMAIN 826 853 CX.	
FT	DOMAIN 867 1013 VWFP 3.	
FT	DOMAIN 1277 1453 VWFA 1.	
FT	DOMAIN 1498 1665 VWFA 2.	
FT	DOMAIN 1691 1871 VWFA 3.	
FT	DOMAIN 1950 2102 VWFD 4.	
FT	DOMAIN 2216 2261 E2.	
FT	DOMAIN 2326 2336 VWFC 1.	
FT	DOMAIN 2429 2495 VWFC 2.	
FT	DOMAIN 2580 2650 VWFC 3.	
FT	DOMAIN 2724 2812 CTCK	
FT	SITE 531 533 CELL ATTACHMENT SITE (POTENTIAL).	
FT	SITE 531 533 CELL ATTACHMENT SITE (POTENTIAL).	
FT	SITE 5607 2500 CELL ATTACHMENT SITE (POTENTIAL).	
FT	DISULFID 767 808 BY SIMILARITY.	
FT	DISULFID 776 804 BY SIMILARITY.	
FT	DISULFID 810 821 BY SIMILARITY.	
FT	DISULFID 867 906 BY SIMILARITY.	

FT	D1S0L.E1D	884	1041	HY SIMILARITY	..	(POTENTIAL)
FT	D1S0L.E1D	894	993	HY SIMILARITY	..	(POTENTIAL)
FT	D1S0L.E1D	914	921	HY SIMILARITY	..	(POTENTIAL)
FT	D1S0L.E1D	1060	1084	HY SIMILARITY	..	(POTENTIAL)
FT	D1S0L.E1D	1071	1111	HY SIMILARITY	..	(POTENTIAL)
FT	D1S0L.E1D	1089	1091	HY SIMILARITY	..	(POTENTIAL)
FT	D1S0L.E1D	1154	1145	HY SIMILARITY	..	(POTENTIAL)
FT	D1S0L.E1D	1144	1169	HY SIMILARITY	..	(POTENTIAL)
FT	D1S0L.E1D	1126	1140	HY SIMILARITY	..	(POTENTIAL)
FT	D1S0L.E1D	1196	1199	HY SIMILARITY	..	(POTENTIAL)
FT	D1S0L.E1D	1244	1237	HY SIMILARITY	..	(POTENTIAL)
FT	D1S0L.E1D	1272	1458	HY SIMILARITY	..	(POTENTIAL)
FT	D1S0L.E1D	1669	1670	HY SIMILARITY	..	(POTENTIAL)
FT	D1S0L.E1D	1686	1872	HY SIMILARITY	..	(POTENTIAL)
FT	D1S0L.E1D	1879	1904	HY SIMILARITY	..	(POTENTIAL)
FT	D1S0L.E1D	1899	1940	OR 1942 (HY SIMILARITY)	..	(POTENTIAL)
FT	D1S0L.E1D	1972	2123	HY SIMILARITY	..	(POTENTIAL)
FT	D1S0L.E1D	1950	2085	HY SIMILARITY	..	(POTENTIAL)
FT	D1S0L.E1D	1927	2088	HY SIMILARITY	..	(POTENTIAL)
FT	D1S0L.E1D	1994	2001	HY SIMILARITY	..	(POTENTIAL)
FT	D1S0L.E1D	2724	2774	HY SIMILARITY	..	(POTENTIAL)
FT	D1S0L.E1D	2739	2788	HY SIMILARITY	..	(POTENTIAL)
FT	D1S0L.E1D	2750	2804	HY SIMILARITY	..	(POTENTIAL)
FT	D1S0L.E1D	2754	2806	HY SIMILARITY	..	(POTENTIAL)
FT	D1S0L.E1D	?	2811	HY SIMILARITY	..	(POTENTIAL)
FT	CAB0.H1D	99	99	N-LINKED (GLYNAC..)	..	(POTENTIAL)
FT	CAB0.H1D	156	156	N-LINKED (GLYNAC..)	..	(POTENTIAL)
FT	CAB0.H1D	211	211	N-LINKED (GLYNAC..)	..	(POTENTIAL)
FT	CAB0.H1D	666	666	N-LINKED (GLYNAC..)	..	(POTENTIAL)
FT	CAB0.H1D	857	857	N-LINKED (GLYNAC..)	..	(POTENTIAL)
FT	CAB0.H1D	1241	1241	N-LINKED (GLYNAC..)	..	(POTENTIAL)
FT	CAB0.H1D	1515	1515	N-LINKED (GLYNAC..)	..	(POTENTIAL)
FT	CAB0.H1D	1574	1574	N-LINKED (GLYNAC..)	..	(POTENTIAL)
FT	CAB0.H1D	2253	2253	N-LINKED (GLYNAC..)	..	(POTENTIAL)
FT	CAB0.H1D	2290	2290	N-LINKED (GLYNAC..)	..	(POTENTIAL)
FT	CAB0.H1D	2457	2457	N-LINKED (GLYNAC..)	..	(POTENTIAL)
FT	CAB0.H1D	2400	2400	N-LINKED (GLYNAC..)	..	(POTENTIAL)
FT	CAB0.H1D	2546	2546	N-LINKED (GLYNAC..)	..	(POTENTIAL)
FT	CAB0.H1D	2585	2585	N-LINKED (GLYNAC..)	..	(POTENTIAL)
FT	CAB0.H1D	2790	2790	N-LINKED (GLYNAC..)	..	(POTENTIAL)
FT	CAB0.H1D	95	95	G -> E (IN REF. 2)	..	(POTENTIAL)
FT	CAB0.H1D	70	70	V -> I (IN REF. 2)	..	(POTENTIAL)
FT	CAB0.H1D	266	266	A -> C (IN REF. 2)	..	(POTENTIAL)
FT	CAB0.H1D	280	280	I -> V (IN REF. 2)	..	(POTENTIAL)
FT	CAB0.H1D	409	411	WCH -> ICG (IN REF. 2)	..	(POTENTIAL)
FT	CAB0.H1D	994	994	G -> A (IN REF. 1)	..	(POTENTIAL)
FT	CAB0.H1D	1021	1021	E -> L (IN REF. 1)	..	(POTENTIAL)
FT	CAB0.H1D	2381	2381	L -> P (IN REF. 2)	..	(POTENTIAL)
FT	CAB0.H1D	2406	2406	P -> I (IN REF. 2)	..	(POTENTIAL)
FT	CAB0.H1D	2813 AA	309716 MW	50F93E1F5E72F80C CRC64	..	(POTENTIAL)
QY	4 MCFEENKPVSSKSEKSTKRLKLVETLHKKLVNVLVYVCHGVVPEVPEYKNDK 53	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100	
DB	650 LSTQWQVYVYGGTPTPTNMTKSLSTPEPKNE-----VLLKAPCPHRLALDEROD 700	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62	
QY	64 CVSAEDP 70	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100	
DB	701 CVPKAC 707	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62	
RESULT	6					
ANALYSE	ANALYSE	STANDARD	PRG	115 AA		
AN	01-0071-2000 (Ref. 40, last sequence update)					
DB	01-0071-2000 (Ref. 40, last sequence update)					
DB	01-0071-2000 (Ref. 40, last annotation update)					
DB	ACCESSORY GLAND PROTEIN AC6262					

OS ACPEP-6-G01362.
CS trophoblast melanogaster (fruit fly).
CC Eubryonia; Metazoa; Anthropoda; Tracheata; Insecta;
OC Ephyridia; Neoptera; Endopterygota; Diptera; brachytera; Platycomorpha;
OC Ephyrididae; Trophoblastidae; trophoblasta.
XX NCBI_taxid=7227?
OT
RP
RP SEQUENCE FROM N.A.
RC STRAIN CANTON-S; TISSUE Male accessory gland;
RX MEDLINE:20196005; PubMed:10731132.
RA Adams M.D., Colniker S.E., Holt R.A., Evans G.A., Goodyear J.D.,
RA Adamalides P.G., Scherer S.E., Li P.W., Hoskins R.A., Gallo R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Y., Chen L.X.,
RA Brandon R.C., Rogers Y.H.C., Blazer K.G., Champo M., Pfeiffer B.D.,
RA Wae K.L., Boyle A.C., Baxter E.G., Heit G., Nelson C.R., Milos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballou P.M., Basu A., Bayram T., Beyarav A., Bozinger T., Bradbury F.M.,
RA Brown S.Y., Bruns S.V., Buchanan B.P., Chandrasekhar V., Chaturvedi S.,
RA Borokova D., Botchan M.P., Boucek J., Brockstein P., Brotter P.,
RA Burks K.C., Busam D.A., Butler H., Calfee E., Cantor A., Chandler I.,
RA Cherry J.M., Chew G.S., Dalbey C., Danopoulos L.B., Davies P.,
RA de Paulis B., Delcher A., Deng Z., Deodhi A.D., Dew I., Dieckmann S.M.,
RA Dodson K., Dong L.B., Downes M., Duggan Snyder S., Durkin B.G., Dunn P.,
RA Little L.R., Edwards-Rizvi C.E., Feltus F.A., Fertig S., Fleischman W.,
RA Foster C., Gabrielian A.B., Gait N.S., Gilbert W.M., Glasser K.,
RA Glöckner A., Gong F., Gorell T.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.F., Hoek J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ideyan C.,
RA Jarati M., Kaish P., Karpen G.H., Ke Z., Kristianson J.A., Kretzschmar K.A.,
RA Kimmel B.E., Kodali G.D., Kiddle C., Klaczko S., Kulip D., Lal Z.,
RA Liao X., Lee J.Y., Levitsky A.A., Li Q., Li Z., Lin Y., Liu X.,
RA Liu X., Mattei P., McIntosh T.C., McLeod M.P., McPherson D.,
RA Metcalfe C., Mishina N.V., Murray L., Morris J., Mostoslavsky A.,
RA Mount S.M., Moy M., Murphy P., Murphy L., Murty D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusseroth D.P., Pacheco J.M.,
RA Palczowski M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shier B.C., Siden-Kiamos I., Simpson M., Skopetski M.P., Smith T.,
RA Spier F., Spradling A.C., Stapleton M., Strong R., Sun F.,
RA Svirskaas R., Taylor C., Turner R., Venter E., Wang A.B., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Winick-Lock G.M., Wisselbacher J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye S., Yeh K.-F., Yancet J.-S., Zhao M., Zhou X., Zhu S., Zheng H.-Q.,
RA Zhang X.H., Zhang F.N., Zhang W., Zhou X., Zhu S., Zhu X., Smith H.-Q.,
RA Gibbs K.A., Myers E.W., Rubin G.M., Venter J.C.?
RI "The genome sequence of *Trophoblasta melanogaster*".
SI Science 287:2185-2195 (2000).
CT -1- FUNCTION: RESPONSIBLE FOR PHYSIOLOGICAL AND BEHAVIORAL CHANGES IN
CT MALE FEMALE FLIES.
CT -1- SUBCELLULAR LOCATION: SECRETED (POTENTIAL).
CT
CT This SWISS-PROT entry is copyright . It is produced through a collaboration
CT between the Swiss Institute of Bioinformatics and the EMBL collaboration
CT the European Bioinformatics Institute. There are no restrictions on its
CT use by non-profit institutions as long as its content is in no way
CT modified and this statement is not removed, usage by and for commercial
CT entities requires a license agreement (<http://www.isb-sib.ch/about/copyright>
CT or send an email to license@sib-ch).
DR EMBL: U05763; AAB96387.1; --
DR EMBL: AF001475; AAP47683.1; --
DR FlyBase: FBgn0020509; AcGP62P.
RW bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted May 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

```

30 SEQUENCE 115 AA; 12679 MW; 4326AAAEF032D94DFE254;
    QUERY MATCH          17.9% SCORE 87; DB 1; LENGTH 115;
    POST LOCAL SIMILARITY 31.0%; PRED NO 0.042;
    MATCHES      22; CONSERVATIVE 8; MISMATCHES 23; GAPS 12; GAPS 4
OY 4 MOONENKYSTOSIS-RELATED VITREOPATHIC PROTEIN [WU] (VF004)
DB   32 VF004-1-IGTFT PVN-PELVICYS-----NM--VM--APAV--PQVYINERIP 60
OY 63 KVSADPCELD 73
EB   81 ACVLSPDCPKD 91
RESULT 7
LBP2_HUMAN STANDARD; PRI: 4655 AA.
AC p98164; 000711; Q16215;
PI 01-OCT-1996 (Rel. 34; Created)
DI 15-DEC-1998 (Rel. 37; Last sequence update)
DI 01-OCT-2000 (Rel. 40; Last annotation update)
DE LOW-DENSITY LIPOPROTEIN RECEPTOR-RELATED PROTEIN 2 PRECURSOR (MPTANIN)
PF (Glycoprotein 330)
GN LRP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Primates; Primates; Catarrhini; Hominoidea; Homin;
OX NCBI_TaxID=9606;
RN 111
RP SEQUENCE FROM N A
RC TISSUE=Kidney;
RA MEDLINE=96395376; PUBMED 8706697;
RA Hjaelm G., Murray F., Cromley C., Harazin M., Lindgren S., Gragstad L.,
RA EK R., Larsson M., Juhlin C., Hellman P., Davis B., Akerstrom G.,
RA Kask L., Morse B.;
RT "Cloning and sequencing of human gp330, a (alpha?)2-binding receptor
R1 with potential intracellular signaling properties."
R1 Eur. J. Biochem. 239:132-137(1996).
RN 121
RP SEQUENCE OF 2705-4453 FROM N.A.
RC TISSUE=Kidney;
RA Knauk C., Argraves W.S.;
RI Submitted (DEC-1993) TO THE EMBL/GenBank/DDBJ databases.
RN 131
RP SEQUENCE OF 3433-4453 FROM N A
RC TISSUE=Kidney;
RA MEDLINE=95048337, PubMed-759375;
RA Korenberg J.R., Argraves K.M., Chen X.N., Tran H.;
RA Strickland D.K., Argraves W.S.;
RT "Chromosomal localization of human genes for the LDL receptor family
R1 member q1q24-q1q21-30 (1412) and its associated protein RAH
R1 (LRPAI)."
RL Genomics 22:88-93(1994).
RN 141
RP SEQUENCE OF 4139-4406 FROM N.A.
RA MEDLINE=94244704; PUBMED 8187828;
RA Lindgren S., Hjaelm G., Hellman P., EK R., Juhlin C., Rastad J.,
RA Ptaretsky I., Akerstrom G., Park L.;
RT "A protein involved in calcium sensing of the human parathyroid and
R1 placental cytotrophoblast cells belongs to the LDL-receptor protein
R1 superfamily."
RL Exp. Cell Res. 212:344-350(1994).
RN 151
RP FUNCTION.
RA MEDLINE=95286588; PubMed 7748901;
RA Komoroski M.Z., Louikova P.B., Stefanescu S., Hartung T.A.F.,
RA Brower R.H., Strickland D.K., Argraves W.S.;
RT "Identification of glycoprotein 330 as an embryonic receptor for
R1 apolipoprotein 17/18."
RI Biol. Chem. 270:13070-13075(1995).
RI -1- FUNCTION: BINDS SPECIFICALLY CLUSTERS WITH HIGH AFFINITY, NOT

```

CC	ALSTILANS IN COMMON WITH OTHER FAMILY MEMBERS: PLASMINOGEN,
CC	EXTRACELLULAR MATRIX COMPONENTS, PLASMINOGEN ACTIVATOR-PLASMINOGEN
CC	ACTIVATOR INHIBITOR TYPE 1 COMPLEX, APOLIPOPROTEIN E-ENRICHED
CC	BETA-VLDL, LIPOPROTEIN LIPASE, LACTOFERRIN AND CALCIUM.
CC	-1- FUNCTION: PROPEPTIDE-MEDIATED HYDRAE OF POLYMERIC DRUGS SUCH AS
CC	ACELININ, AMINOLIPIDATES AND POLYMERS B (4% SIMILARITY).
CC	-1- FUNCTION: MAY PARTICIPATE IN REGULATION OF PARATHYROID-HORMONE AND
CC	PABA-HYDROLYTIC-HORMONE-RELATED PROTEIN RELEASE.
CC	-1- SUBUNIT: FORMS A MULTIMERIC COMPLEX TOGETHER WITH A RECEPTOR-
CC	ASSOCIABLE PROTEIN (PAB).
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN,
CC	TISSUE SPECIFICITY: ABSORPTIVE EPITHELIA, INCLUDING RENAL
CC	PROXIMAL TUBULES.
CC	-1- SIMILARITY: CONTAINS 36 LDL-RECEPTOR CLASS A DOMAINS.
CC	-1- SIMILARITY: CONTAINS 37 LDL-RECEPTOR CLASS B DOMAINS.
CC	-1- SIMILARITY: CONTAINS 4 SH3-BINDING DOMAINS.
CC	-1- SIMILARITY: CONTAINS 1 SH2-BINDING DOMAIN.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation at
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed, changed by and for commercial
CC	or other purposes. If you wish to quote or send an email to license@swissprot.ch).
CC	-----
DN	FMN1; H33837; AAAH1649.1; -
DN	FMN1; H04441; AAH02892.1; -
DN	FMN1; S73145; AAH02825.1; -
DN	HSSP; P01130; JAL1.
DN	MIM; 600073; -
DR	InterPro: IPR000033; -
DR	InterPro: IPR000152; -
DR	InterPro: IPR000561; -
DR	InterPro: IPR001881; -
DR	InterPro: IPR002172; -
DP	Pfam; PF00904; FGF_13.
DR	Pfam; PF00057; ldl_recept_a; 36.
DR	Pfam; PF00058; ldl_recept_b; 37.
DR	PRINTS; PR00261; LDLRECEPTOR.
DR	PROSITE; PS00010; ASX_HYDROXYL_4.
DR	PROSITE; PS00032; EGF_1; 1.
DR	PROSITE; PS01186; EGF_2; 9.
DR	PROSITE; PS01187; EGF_CA; 3.
DR	PROSITE; PS01209; LDLR_A_1; 31.
DR	PROSITE; PS00068; LDLR_A_2; 36.
KW	Glycoprotein, Fc-gamma, Endocytosis, Cyt-d fil., Transmembrane,
KW	Receptor, Egg-like domain, Signal, Polymorphism.
FT	SIGNAL 1 25
FT	CDSIN 26 4655
FT	-----
FT	L-DOMAIN 25 4473
FT	TRANSMEM 4424 4446
FT	DOMAIN 4447 4655
FT	DOMAIN 26 64
FT	DOMAIN 65 105
FT	DOMAIN 106 144
FT	DOMAIN 145 181
FT	DOMAIN 182 219
FT	DOMAIN 220 258
FT	DOMAIN 264 308
FT	DOMAIN 309 347
FT	DOMAIN 348 386
FT	DOMAIN 436 477
FT	DOMAIN 479 520
FT	DOMAIN 522 567
FT	DOMAIN 569 613
FT	DOMAIN 613 653
FT	DOMAIN 654 705
FT	DOMAIN 753 794
FT	DOMAIN 796 836
FT	DOMAIN 838 880
FT	-----
FT	LOW-INTENSITY LIPOPROTEIN RECEPTOR-RELATED
FT	PROTEIN 2.
FT	EXTRAECYLLIMAR (POTENTIAL).
FT	POTENTIAL.
FT	CYTOSOLASMIC (POTENTIAL).
FT	LDL-RECEPTOR CLASS A 1.
FT	LDL-RECEPTOR CLASS A 2.
FT	LDL-RECEPTOR CLASS A 3.
FT	LDL-RECEPTOR CLASS A 4.
FT	LDL-RECEPTOR CLASS A 5.
FT	LDL-RECEPTOR CLASS A 6.
FT	LDL-RECEPTOR CLASS A 7.
FT	EGF-LIKE 1.
FT	EGF-LIKE 2.
FT	LDL-RECEPTOR CLASS B 1.
FT	LDL-RECEPTOR CLASS B 2.
FT	LDL-RECEPTOR CLASS B 3.
FT	LDL-RECEPTOR CLASS B 4.
FT	LDL-RECEPTOR CLASS B 5.
FT	EGF-LIKE 3.
FT	LDL-RECEPTOR CLASS B 6.
FT	LDL-RECEPTOR CLASS B 7.
FT	LDL-RECEPTOR CLASS B 8.

FT	DOMAIN	882	924	1014	LDL-RECEPTOR CLASS B 9.	EGF-LIKE 4.	LDL-RECEPTOR CLASS A 8.	LDL-RECEPTOR CLASS A 9.	LDL-RECEPTOR CLASS A 10.	LDL-RECEPTOR CLASS A 11.	LDL-RECEPTOR CLASS A 12.	LDL-RECEPTOR CLASS A 13.	LDL-RECEPTOR CLASS A 14.	LDL-RECEPTOR CLASS A 15.	EGF-LIKE 5.	EGF-LIKE 6.	LDL-RECEPTOR CLASS B 10.	LDL-RECEPTOR CLASS B 11.	LDL-RECEPTOR CLASS B 12.	LDL-RECEPTOR CLASS B 13.	LDL-RECEPTOR CLASS B 14.	LDL-RECEPTOR CLASS B 15.	LDL-RECEPTOR CLASS B 16.	LDL-RECEPTOR CLASS B 17.	LDL-RECEPTOR CLASS B 18.	LDL-RECEPTOR CLASS B 19.	EGF-LIKE 8.	LDL-RECEPTOR CLASS B 20.	LDL-RECEPTOR CLASS B 21.	LDL-RECEPTOR CLASS B 22.	LDL-RECEPTOR CLASS B 23.	LDL-RECEPTOR CLASS B 24.	EGF-LIKE 9.	LDL-RECEPTOR CLASS B 25.	LDL-RECEPTOR CLASS B 26.	LDL-RECEPTOR CLASS B 27.	LDL-RECEPTOR CLASS B 28.	LDL-RECEPTOR CLASS B 29.	EGF-LIKE 10.	LDL-RECEPTOR CLASS A 16.	LDL-RECEPTOR CLASS A 17.	LDL-RECEPTOR CLASS A 18.	LDL-RECEPTOR CLASS A 19.	LDL-RECEPTOR CLASS A 20.	LDL-RECEPTOR CLASS A 21.	LDL-RECEPTOR CLASS A 22.	LDL-RECEPTOR CLASS A 23.	LDL-RECEPTOR CLASS A 24.	LDL-RECEPTOR CLASS A 25.	EGF-LIKE 11.	EGF-LIKE 12.	CALCIUM BINDING (POTENTIAL).	LDL-RECEPTOR CLASS B 30.	LDL-RECEPTOR CLASS B 31.	LDL-RECEPTOR CLASS B 32.	LDL-RECEPTOR CLASS B 33.	LDL-RECEPTOR CLASS B 34.	EGF-LIKE 13.	LDL-RECEPTOR CLASS A 26.	LDL-RECEPTOR CLASS A 27.	LDL-RECEPTOR CLASS A 28.	LDL-RECEPTOR CLASS A 29.	LDL-RECEPTOR CLASS A 30.	LDL-RECEPTOR CLASS A 31.	LDL-RECEPTOR CLASS A 32.	LDL-RECEPTOR CLASS A 33.	LDL-RECEPTOR CLASS A 34.	LDL-RECEPTOR CLASS A 35.	LDL-RECEPTOR CLASS A 36.	EGF-LIKE 14.	EGF-LIKE 15.	CALCIUM-BINDING (POTENTIAL).	LDL-RECEPTOR CLASS B 35.	LDL-RECEPTOR CLASS B 36.	LDL-RECEPTOR CLASS B 37.	EGF-LIKE 16.	EGF-LIKE 17.						
FT	DOMAIN	1024	1062	1103	1137	1145	1185	1224	1268	1269	1307	1310	1350	1389	1429	1478	1519	1562	1608	1653	1695	1730	1741	1841	1843	1881	1929	1971	2012	2059	2157	2195	2250	2292	2344	2388	2431	2476	2517	2561	2603	2604	2646	2693	2738	2777	2819	2861	2901	2945	2990	3029	3070	3110	3151	3192	3230	3282	3331	3375	3418	3459	3495	3509	3550	3591	3632	3673	3716	3756	3795	3834	3880	3922	3964	4006	4048	4195	4239	4282	4330	4377	4411
FT	DOMAIN	1024	1062	1103	1137	1145	1185	1224	1268	1269	1307	1310	1350	1389	1429	1478	1519	1562	1608	1653	1695	1730	1741	1841	1843	1881	1929	1971	2012	2059	2157	2195	2250	2292	2344	2388	2431	2476	2517	2561	2603	2604	2646	2693	2738	2777	2819	2861	2901	2945	2990	3029	3070	3110	3151	3192	3230	3282	3331	3375	3418	3459	3495	3509	3550	3591	3632	3673	3716	3756	3795	3834	3880	3922	3964	4006	4048	4195	4239	4282	4330	4377	4411

[illegible]

DE (VITELLOGENIN RECEPTOR) (VTG RECEPTOR).
 GN VITR OR VTGR.
 OS Gallus gallus (Chicken).
 OC Pharyngodon, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX MBLI_TAXID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 PC STEIN-WHITE LECTURE: TISSUE-SPECIFIC;
 EX MEDLINE=95045409; PubMed=7957081;
 RA Buto H., Hermann M., Nadelin M.O., Jacobsen L., Steward S.,
 RA Nimpf J., Yamamoto T., Schneider W.J.;
 RT "Chicken vitellogenin is mediated by an eight lipid-binding repeat
 member of the LDL receptor family.";
 RL EMBO J. 13:5165-5175(1994).
 RN [2]
 RP SEQUENCE OF 510-518: 546-554 AND 819-827.
 PC STEIN-WHITE LECTURE: TISSUE-POLYMER membrane;
 EX MEDLINE=9201638; PubMed 1655760;
 RA Barber D.L., Sanders E.J., Abersold R., Schneider W.J.;
 RA Barber D.L., Sanders E.J., Abersold R., Schneider W.J.;
 RT "The receptor for yolk lipoprotein deposition is a 10-subunit complex."
 OC 1. FUNCTION: BINDS VITR AND VLDL TRANSPORTS THEM INTO CELLS BY
 CC ENDOCYTOSIS. IN ORDER TO BE INTERNALIZED, THE RECEPTOR LIGAND
 CC COMPLEXES MUST FIRST CLUSTER INTO CLATHRIN-COATED PITS.
 CC 1. TISSUE SPECIFICITY: ADULTANT IN YOLK; MUCH LESS IN HEART AND
 CC SKELETAL MUSCLE.
 CC 1. SIMILARITY: CONTAINS 8 LDL-RECEPTOR CLASS A DOMAINS.
 CC 1. SIMILARITY: CONTAINS 6 LDL-RECEPTOR CLASS B DOMAINS.
 CC 1. SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/seqdb/>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR PF00107: CAAS6505.1;
 DR HSSP: P01130; IAJJ.
 DR InterPro: IPR000033;
 DR InterPro: IPR000152;
 DR InterPro: IPR000561;
 DR InterPro: IPR001881;
 DR InterPro: IPR002172;
 DR Pfam: PF00057; Idr_recept_a; 8.
 DR Pfam: PF00058; Idr_recept_b; 5.
 DR PRINTS: PR00261; IdrRecept.
 DR PROSITE: PS00010; ASX_HYPOXYL; 2.
 DR PROSITE: PS00022; EGF_1; FALSE_NGCG.
 DR PROSITE: PS01186; EGF_2; 3.
 DR PROSITE: PS01187; EGF_CA; 2.
 DR PROSITE: PS01209; IdrR_A; 2.
 DR PROSITE: PS00068; IdrR_2; 8.
 DR Glycoprotein, VLDL, cholesterol metabolism, lipid transport;
 KM Endocytosis, coated pits, transmembrane receptor, signal;
 KM EGF-like domain; Repeat.
 FT SIGNAL 1 43
 FT CHAIN 44 863
 FT DOMAIN 44 785
 FT TRANSMEM 786 809
 FT DOMAIN 810 863
 FT DOMAIN 49 87
 FT DOMAIN 88 128
 FT DOMAIN 129 169
 FT DOMAIN 170 208
 FT DOMAIN 209 249
 FT DOMAIN 255 293
 FT DOMAIN 294 332
 FT DOMAIN 334 373
 FT POTENTIAL.
 FT VERY LOW-DENSITY LIPOPROTEIN RECEPTOR.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT LDL-RECEPTOR CLASS A 1.
 FT LDL-RECEPTOR CLASS A 2.
 FT LDL-RECEPTOR CLASS A 3.
 FT LDL-RECEPTOR CLASS A 4.
 FT LDL-RECEPTOR CLASS A 5.
 FT LDL-RECEPTOR CLASS A 6.
 FT LDL-RECEPTOR CLASS A 7.
 FT LDL-RECEPTOR CLASS A 8.

FT DOMAIN 374 413
 FT DOMAIN 414 453
 FT REPEAT 457 498
 FT REPEAT 499 544
 FT REPEAT 545 587
 FT REPEAT 588 631
 FT REPEAT 632 674
 FT REPEAT 675 716
 FT REPEAT 717 757
 FT DOMAIN 758 827
 FT DOMAIN 828 863
 FT DISULFID 51 63
 FT DISULFID 58 76
 FT DISULFID 70 85
 FT DISULFID 97 115
 FT DISULFID 109 126
 FT DISULFID 131 145
 FT DISULFID 138 158
 FT DISULFID 152 167
 FT DISULFID 172 184
 FT DISULFID 179 197
 FT DISULFID 191 206
 FT DISULFID 211 223
 FT DISULFID 232 236
 FT DISULFID 230 247
 FT DISULFID 257 269
 FT DISULFID 264 282
 FT DISULFID 276 291
 FT DISULFID 296 308
 FT DISULFID 303 321
 FT DISULFID 315 330
 FT DISULFID 336 349
 FT DISULFID 344 362
 FT DISULFID 356 373
 FT DISULFID 378 389
 FT DISULFID 385 398
 FT DISULFID 400 412
 FT DISULFID 418 428
 FT DISULFID 424 437
 FT DISULFID 439 452
 FT DISULFID 726 734
 FT DISULFID 735 754
 FT DISULFID 756 769
 FT CARBOHYD 169 199
 FT CARBOHYD 773 773
 FT SEQUENCE 653 AA; 94704 MW; 067280748FA2245 CRC64;
 Query Match 17.6%; Score 85.5; DB 1; Length 863;
 Best Local Similarity 33.3%; Pred. No. 0.3;
 Matches 25; Conservative 9; Mismatches 36; Indels 5; Gaps 4;
 27 4 MDCSENEKVEDSUSKEC-IRKAKTVVEELDEPNVPCVVPVCHGVV-CPKPTPKND 61
 28 255 VQSTSEV-QQSGSPCHKKWPKVWPKVQKQSPQNPSPPTPPVQVPCVPCVPCVPCV 312
 29 62 DKCVASPCPLDND 76
 30 313 KQTNVVRK-LEGD 326
 RESULT 10
 LTVR_HUMAN
 ID LTVR_HUMAN STANDARD; PRT; 873 AA.
 AC P98155;
 DT 01-OCT-1996 (Ref. 34, Created)
 DT 01-OCT-1996 (Ref. 34, Last annotation update)
 DT 01-OCT-2000 (Ref. 40, Last sequence update)
 DE VERY LOW-DENSITY LIPOPROTEIN RECEPTOR (VLDL RECEPTOR).
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

00 NBI_FaxID 9605;
 01 [1]
 02 SEQUENCE FROM N.A.
 03 TISSUE: Skeletal muscle;
 04 MEDLINE 9417478; PubMed 812815;
 05 Givovels M.E., Caird M., Britz D., Jackson C.I., Patterson D.,
 06 Strauss J.F.?
 07 "Cloning of a cDNA encoding a putative human very low density
 08 lipoprotein/apoE receptor and assignment of the gene to
 09 chromosome 9p11-p23."
 10 Genet. Cell Mol. Genet. 19:557-569(1993).
 11 [2]
 12 SEQUENCE FROM N.A.
 13 TISSUE: Heart;
 14 MEDLINE 94348496; PubMed 8069294;
 15 Webb J.C., Patel D.D., Jones M.D., Knight R.L., Soutar A.K.?
 16 "Characterization and tissue-specific expression of the human 'very
 17 low density lipoprotein (VLDL) receptor' mRNA."
 18 Hum. Mol. Genet. 3:531-537(1994).
 19 [3]
 20 SEQUENCE FROM N.A.
 21 MEDLINE 94124575; PubMed 8294473;
 22 Sakai I., Hashino A., Takahashi S., Miura Y., Ishii H., Suzuki H.,
 23 Kawarabayashi Y., Yamamoto T.?
 24 "Structure, chromosome location, and expression of the human very low
 25 density lipoprotein receptor gene."
 26 J. Biol. Chem. 269:2173-2182(1994).
 27 [4]
 28 SEQUENCE FROM N.A.
 29 TISSUE: Heart;
 30 MEDLINE 94292216; PubMed 8020981;
 31 Oka K., Fain G.W., Sullivan M., Lindsay E., Baldini A., Chan L.?
 32 "Human very low density lipoprotein receptor complementary DNA and
 33 deduced amino acid sequence and localization of its gene (VLDLR) to
 34 chromosome band 9p24 by fluorescence in situ hybridization."
 35 Genomics 20:298-300(1994).
 36 [5]
 37 FUNCTION: BINDS VLDL AND TRANSPORTS IT INTO CELLS BY ENDOCYTOSIS.
 38 IN ORDER TO BE INTERNALIZED, THE RECEPTOR-LIGAND COMPLEXES MUST
 39 FIRST CLUSTER INTO CLATHRIN-COATED PITS (BY SIMILARITY).
 40 [6]
 41 SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 42 [7]
 43 TISSUE SPECIFICITY: ABUNDANT IN HEART AND SKELETAL MUSCLE; ALSO
 44 OVARY AND KIDNEY; NOT IN LIVER.
 45 [8]
 46 SIMILARITY: CONTAINS 8 LDL-RECEPTOR CLASS A DOMAINS.
 47 [9]
 48 SIMILARITY: CONTAINS 6 LDL-RECEPTOR CLASS B DOMAINS.
 49 [10]
 50 SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
 51 [11]
 52 This Swiss-Prot entry is copyright. It is produced through a collaboration
 53 between the Swiss Institute of Bioinformatics and the EMBL outstation
 54 at the European Bioinformatics Institute. There are no restrictions on its
 55 use by non-profit institutions as long as its content is in no way
 56 modified and this statement is not removed. Usage by and for commercial
 57 entities requires a license agreement (See <http://www.isb-sib.ch/annotation/>
 58 or send an email to license@isb-sib.ch).
 59 -----
 60 EMBL: L20470; AAA53684.1; -;
 61 EMBL: D165312; BAA03969.1; -;
 62 EMBL: D164495; BAA03969.1; JOINED.
 63 EMBL: D16508; BAA03969.1; JOINED.
 64 EMBL: D16510; BAA03969.1; JOINED.
 65 EMBL: D16514; BAA03969.1; JOINED.
 66 EMBL: D16516; BAA03969.1; JOINED.
 67 EMBL: D16518; BAA03969.1; JOINED.
 68 EMBL: D16520; BAA03969.1; JOINED.
 69 EMBL: D16522; BAA03969.1; JOINED.
 70 EMBL: D16523; BAA03969.1; JOINED.
 71 EMBL: D16524; BAA03969.1; JOINED.
 72 EMBL: D16525; BAA03969.1; JOINED.
 73 EMBL: D16526; BAA03969.1; JOINED.
 74 EMBL: D16527; BAA03969.1; JOINED.
 75 EMBL: D16528; BAA03969.1; JOINED.
 76 EMBL: D16529; BAA03969.1; JOINED.
 77 EMBL: D16530; BAA03969.1; JOINED.
 78 EMBL: D16531; BAA03969.1; JOINED.

DR EMBL: S73849; AAB31735.1; -;
 DR EMBL: D16493; BAA03945.1; -;
 DR EMBL: D16494; BAA03946.1; -;
 DR EMBL: L22431; AAA61344.1; -;
 DR HSSP: P01130; IAIJ.
 DR MIM: 192977; -;
 DR InterPro: IPR000033; -;
 DR InterPro: IPR000152; -;
 DR InterPro: IPR000561; -;
 DR InterPro: IPR001861; -;
 DR InterPro: IPR02172; -;
 DR Pfam: PF00057; 1d1_recept_a; 8.
 DR Pfam: PF00058; 1d1_recept_b; 5.
 DR PRINTS: PR00261; LDLRECEPTOR.
 DR PROSITE: PS00010; ASX_HYDROXYL; 2.
 DR PROSITE: PS00022; EGF_1; FALSE_NEG.
 DR PROSITE: PS01185; EGF_2; 3.
 DR PROSITE: PS01187; EGF_CA; 1.
 DR PROSITE: PS01209; LDLRA_1; 8.
 DR PROSITE: PS00068; LDLRA_2; 8.
 DR GlycoProtex: VLDL; Cholesterol metabolism; Lipid transport;
 DR EndoCytoSis: Coated pits; Transmembrane receptor; Signal;
 DR EGF_Like_Domain; Repeat; Attractive signal.
 DR EFT SIGNAL 1 27
 DR CHAIN 28 873
 DR DOMAIN 28 797
 DR IRANSEMEM 798 819
 DR DOMAIN 820 873
 DR DOMAIN 31 69
 DR DOMAIN 70 110
 DR DOMAIN 111 151
 DR DOMAIN 152 190
 DR DOMAIN 191 231
 DR DOMAIN 237 275
 DR DOMAIN 276 314
 DR DOMAIN 316 355
 DR DOMAIN 356 395
 DR DOMAIN 396 435
 DR REPEAT 439 480
 DR REPEAT 481 524
 DR REPEAT 525 567
 DR REPEAT 568 611
 DR REPEAT 612 654
 DR REPEAT 655 696
 DR DOMAIN 702 750
 DR DOMAIN 751 790
 DR DOMAIN 832 837
 DR DISULFID 33 45
 DR DISULFID 40 58
 DR DISULFID 52 67
 DR DISULFID 72 84
 DR DISULFID 79 97
 DR DISULFID 91 108
 DR DISULFID 113 127
 DR DISULFID 120 140
 DR DISULFID 134 149
 DR DISULFID 154 166
 DR DISULFID 161 179
 DR DISULFID 173 188
 DR DISULFID 193 205
 DR DISULFID 200 218
 DR DISULFID 212 229
 DR DISULFID 239 251
 DR DISULFID 251 264
 DR DISULFID 268 273
 DR DISULFID 278 290
 DR DISULFID 285 303
 DR DISULFID 297 312
 DR DISULFID 318 331
 DR DISULFID 326 344
 DR DISULFID 338 355
 DR DISULFID 360 371
 DR DOMAIN 832 837
 DR DISULFID 33 45
 DR DISULFID 40 58
 DR DISULFID 52 67
 DR DISULFID 72 84
 DR DISULFID 79 97
 DR DISULFID 91 108
 DR DISULFID 113 127
 DR DISULFID 120 140
 DR DISULFID 134 149
 DR DISULFID 154 166
 DR DISULFID 161 179
 DR DISULFID 173 188
 DR DISULFID 193 205
 DR DISULFID 200 218
 DR DISULFID 212 229
 DR DISULFID 239 251
 DR DISULFID 251 264
 DR DISULFID 268 273
 DR DISULFID 278 290
 DR DISULFID 285 303
 DR DISULFID 297 312
 DR DISULFID 318 331
 DR DISULFID 326 344
 DR DISULFID 338 355
 DR DISULFID 360 371
 DR DOMAIN 832 837
 DR DISULFID 33 45
 DR DISULFID 40 58
 DR DISULFID 52 67
 DR DISULFID 72 84
 DR DISULFID 79 97
 DR DISULFID 91 108
 DR DISULFID 113 127
 DR DISULFID 120 140
 DR DISULFID 134 149
 DR DISULFID 154 166
 DR DISULFID 161 179
 DR DISULFID 173 188
 DR DISULFID 193 205
 DR DISULFID 200 218
 DR DISULFID 212 229
 DR DISULFID 239 251
 DR DISULFID 251 264
 DR DISULFID 268 273
 DR DISULFID 278 290
 DR DISULFID 285 303
 DR DISULFID 297 312
 DR DISULFID 318 331
 DR DISULFID 326 344
 DR DISULFID 338 355
 DR DISULFID 360 371
 DR DOMAIN 832 837
 DR DISULFID 33 45
 DR DISULFID 40 58
 DR DISULFID 52 67
 DR DISULFID 72 84
 DR DISULFID 79 97
 DR DISULFID 91 108
 DR DISULFID 113 127
 DR DISULFID 120 140
 DR DISULFID 134 149
 DR DISULFID 154 166
 DR DISULFID 161 179
 DR DISULFID 173 188
 DR DISULFID 193 205
 DR DISULFID 200 218
 DR DISULFID 212 229
 DR DISULFID 239 251
 DR DISULFID 251 264
 DR DISULFID 268 273
 DR DISULFID 278 290
 DR DISULFID 285 303
 DR DISULFID 297 312
 DR DISULFID 318 331
 DR DISULFID 326 344
 DR DISULFID 338 355
 DR DISULFID 360 371
 DR DOMAIN 832 837
 DR DISULFID 33 45
 DR DISULFID 40 58
 DR DISULFID 52 67
 DR DISULFID 72 84
 DR DISULFID 79 97
 DR DISULFID 91 108
 DR DISULFID 113 127
 DR DISULFID 120 140
 DR DISULFID 134 149
 DR DISULFID 154 166
 DR DISULFID 161 179
 DR DISULFID 173 188
 DR DISULFID 193 205
 DR DISULFID 200 218
 DR DISULFID 212 229
 DR DISULFID 239 251
 DR DISULFID 251 264
 DR DISULFID 268 273
 DR DISULFID 278 290
 DR DISULFID 285 303
 DR DISULFID 297 312
 DR DISULFID 318 331
 DR DISULFID 326 344
 DR DISULFID 338 355
 DR DISULFID 360 371
 DR DOMAIN 832 837
 DR DISULFID 33 45
 DR DISULFID 40 58
 DR DISULFID 52 67
 DR DISULFID 72 84
 DR DISULFID 79 97
 DR DISULFID 91 108
 DR DISULFID 113 127
 DR DISULFID 120 140
 DR DISULFID 134 149
 DR DISULFID 154 166
 DR DISULFID 161 179
 DR DISULFID 173 188
 DR DISULFID 193 205
 DR DISULFID 200 218
 DR DISULFID 212 229
 DR DISULFID 239 251
 DR DISULFID 251 264
 DR DISULFID 268 273
 DR DISULFID 278 290
 DR DISULFID 285 303
 DR DISULFID 297 312
 DR DISULFID 318 331
 DR DISULFID 326 344
 DR DISULFID 338 355
 DR DISULFID 360 371
 DR DOMAIN 832 837
 DR DISULFID 33 45
 DR DISULFID 40 58
 DR DISULFID 52 67
 DR DISULFID 72 84
 DR DISULFID 79 97
 DR DISULFID 91 108
 DR DISULFID 113 127
 DR DISULFID 120 140
 DR DISULFID 134 149
 DR DISULFID 154 166
 DR DISULFID 161 179
 DR DISULFID 173 188
 DR DISULFID 193 205
 DR DISULFID 200 218
 DR DISULFID 212 229
 DR DISULFID 239 251
 DR DISULFID 251 264
 DR DISULFID 268 273
 DR DISULFID 278 290
 DR DISULFID 285 303
 DR DISULFID 297 312
 DR DISULFID 318 331
 DR DISULFID 326 344
 DR DISULFID 338 355
 DR DISULFID 360 371
 DR DOMAIN 832 837
 DR DISULFID 33 45
 DR DISULFID 40 58
 DR DISULFID 52 67
 DR DISULFID 72 84
 DR DISULFID 79 97
 DR DISULFID 91 108
 DR DISULFID 113 127
 DR DISULFID 120 140
 DR DISULFID 134 149
 DR DISULFID 154 166
 DR DISULFID 161 179
 DR DISULFID 173 188
 DR DISULFID 193 205
 DR DISULFID 200 218
 DR DISULFID 212 229
 DR DISULFID 239 251
 DR DISULFID 251 264
 DR DISULFID 268 273
 DR DISULFID 278 290
 DR DISULFID 285 303
 DR DISULFID 297 312
 DR DISULFID 318 331
 DR DISULFID 326 344
 DR DISULFID 338 355
 DR DISULFID 360 371
 DR DOMAIN 832 837
 DR DISULFID 33 45
 DR DISULFID 40 58
 DR DISULFID 52 67
 DR DISULFID 72 84
 DR DISULFID 79 97
 DR DISULFID 91 108
 DR DISULFID 113 127
 DR DISULFID 120 140
 DR DISULFID 134 149
 DR DISULFID 154 166
 DR DISULFID 161 179
 DR DISULFID 173 188
 DR DISULFID 193 205
 DR DISULFID 200 218
 DR DISULFID 212 229
 DR DISULFID 239 251
 DR DISULFID 251 264
 DR DISULFID 268 273
 DR DISULFID 278 290
 DR DISULFID 285 303
 DR DISULFID 297 312
 DR DISULFID 318 331
 DR DISULFID 326 344
 DR DISULFID 338 355
 DR DISULFID 360 371
 DR DOMAIN 832 837
 DR DISULFID 33 45
 DR DISULFID 40 58
 DR DISULFID 52 67
 DR DISULFID 72 84
 DR DISULFID 79 97
 DR DISULFID 91 108
 DR DISULFID 113 127
 DR DISULFID 120 140
 DR DISULFID 134 149
 DR DISULFID 154 166
 DR DISULFID 161 179
 DR DISULFID 173 188
 DR DISULFID 193 205
 DR DISULFID 200 218
 DR DISULFID 212 229
 DR DISULFID 239 251
 DR DISULFID 251 264
 DR DISULFID 268 273
 DR DISULFID 278 290
 DR DISULFID 285 303
 DR DISULFID 297 312
 DR DISULFID 318 331
 DR DISULFID 326 344
 DR DISULFID 338 355
 DR DISULFID 360 371
 DR DOMAIN 832 837
 DR DISULFID 33 45
 DR DISULFID 40 58
 DR DISULFID 52 67
 DR DISULFID 72 84
 DR DISULFID 79 97
 DR DISULFID 91 108
 DR DISULFID 113 127
 DR DISULFID 120 140
 DR DISULFID 134 149
 DR DISULFID 154 166
 DR DISULFID 161 179
 DR DISULFID 173 188
 DR DISULFID 193 205
 DR DISULFID 200 218
 DR DISULFID 212 229
 DR DISULFID 239 251
 DR DISULFID 251 264
 DR DISULFID 268 273
 DR DISULFID 278 290
 DR DISULFID 285 303
 DR DISULFID 297 312
 DR DISULFID 318 331
 DR DISULFID 326 344
 DR DISULFID 338 355
 DR DISULFID 360 371
 DR DOMAIN 832 837
 DR DISULFID 33 45
 DR DISULFID 40 58
 DR DISULFID 52 67
 DR DISULFID 72 84
 DR DISULFID 79 97
 DR DISULFID 91 108
 DR DISULFID 113 127
 DR DISULFID 120 140
 DR DISULFID 134 149
 DR DISULFID 154 166
 DR DISULFID 161 179
 DR DISULFID 173 188
 DR DISULFID 193 205
 DR DISULFID 200 218
 DR DISULFID 212 229
 DR DISULFID 239 251
 DR DISULFID 251 264
 DR DISULFID 268 273
 DR DISULFID 278 290
 DR DISULFID 285 303
 DR DISULFID 297 312
 DR DISULFID 318 331
 DR DISULFID 326 344
 DR DISULFID 338 355
 DR DISULFID 360 371
 DR DOMAIN 832 837
 DR DISULFID 33 45
 DR DISULFID 40 58
 DR DISULFID 52 67
 DR DISULFID 72 84
 DR DISULFID 79 97
 DR DISULFID 91 108
 DR DISULFID 113 127
 DR DISULFID 120 140
 DR DISULFID 134 149
 DR DISULFID 154 166
 DR DISULFID 161 179
 DR DISULFID 173 188
 DR DISULFID 193 205
 DR DISULFID 200 218
 DR DISULFID 212 229
 DR DISULFID 239 251
 DR DISULFID 251 264
 DR DISULFID 268 273
 DR DISULFID 278 290
 DR DISULFID 285 303
 DR DISULFID 297 312
 DR DISULFID 318 331
 DR DISULFID 326 344
 DR DISULFID 338 355
 DR DISULFID 360 371
 DR DOMAIN 832 837
 DR DISULFID 33 45
 DR DISULFID 40 58
 DR DISULFID 52 67
 DR DISULFID 72 84
 DR DISULFID 79 97
 DR DISULFID 91 108
 DR DISULFID 113 127
 DR DISULFID 120 140
 DR DISULFID 134 149
 DR DISULFID 154 166
 DR DISULFID 161 179
 DR DISULFID 173 188
 DR DISULFID 193 205
 DR DISULFID 200 218
 DR DISULFID 212 229
 DR DISULFID 239 251
 DR DISULFID 251 264
 DR DISULFID 268 273
 DR DISULFID 278 290
 DR DISULFID 285 303
 DR DISULFID 297 312
 DR DISULFID 318 331
 DR DISULFID 326 344
 DR DISULFID 338 355
 DR DISULFID 360 371
 DR DOMAIN 832 837
 DR DISULFID 33 45
 DR DISULFID 40 58
 DR DISULFID 52 67
 DR DISULFID 72 84
 DR DISULFID 79 97
 DR DISULFID 91 108
 DR DISULFID 113 127
 DR DISULFID 120 140
 DR DISULFID 134 149
 DR DISULFID 154 166
 DR DISULFID 161 179
 DR DISULFID 173 188
 DR DISULFID 193 205
 DR DISULFID 200 218
 DR DISULFID 212 229
 DR DISULFID 239 251
 DR DISULFID 251 264
 DR DISULFID 268 273
 DR DISULFID 278 290
 DR DISULFID 285 303
 DR DISULFID 297 312
 DR DISULFID 318 331
 DR DISULFID 326 344
 DR DISULFID 338 355
 DR DISULFID 360 371
 DR DOMAIN 832 837
 DR DISULFID 33 45
 DR DISULFID 40 58
 DR DISULFID 52 67
 DR DISULFID 72 84
 DR DISULFID 79 97
 DR DISULFID 91 108
 DR DISULFID 113 127
 DR DISULFID 120 140
 DR DISULFID 134 149
 DR DISULFID 154 166
 DR DISULFID 161 179
 DR DISULFID 173 188
 DR DISULFID 193 205
 DR DISULFID 200 218
 DR DISULFID 212 229
 DR DISULFID 239 251
 DR DISULFID 251 264
 DR DISULFID 268 273
 DR DISULFID 278 290
 DR DISULFID 285 303
 DR DISULFID 297 312
 DR DISULFID 318 331
 DR DISULFID 326 344
 DR DISULFID 338 355
 DR DISULFID 360 371
 DR DOMAIN 832 837
 DR DISULFID 33 45
 DR DISULFID 40 58
 DR DISULFID 52 67
 DR DISULFID 72 84
 DR DISULFID 79 97
 DR DISULFID 91 108
 DR DISULFID 113 127
 DR DISULFID 120 140
 DR DISULFID 134 149
 DR DISULFID 154 166
 DR DISULFID 161 179
 DR DISULFID 173 188
 DR DISULFID 193 205
 DR DISULFID 200 218
 DR DISULFID 212 229
 DR DISULFID 239 251
 DR DISULFID 251 264
 DR DISULFID 268 273
 DR DISULFID 278 290
 DR DISULFID 285 303
 DR DISULFID 297 312
 DR DISULFID 318 331
 DR DISULFID 326 344
 DR DISULFID 338 355
 DR DISULFID 360 371
 DR DOMAIN 832 837
 DR DISULFID 33 45
 DR DISULFID 40 58
 DR DISULFID 52 67
 DR DISULFID 72 84
 DR DISULFID 79 97
 DR DISULFID 91 108
 DR DISULFID 113 127
 DR DISULFID 120 140
 DR DISULFID 134 149
 DR DISULFID 154 166
 DR DISULFID 161 179
 DR DISULFID 173 188
 DR DISULFID 193 205
 DR DISULFID 200 218
 DR DISULFID 212 229
 DR DISULFID 239 251
 DR DISULFID 251 264
 DR DISULFID 268 273
 DR DISULFID 278 290
 DR DISULFID 285 303
 DR DISULFID 297 312
 DR DISULFID 318 331
 DR DISULFID 326 344
 DR DISULFID 338 355
 DR DISULFID 360 371
 DR DOMAIN 832 837
 DR DISULFID 33 45
 DR DISULFID 40 58
 DR DISULFID 52 67
 DR DISULFID 72 84
 DR DISULFID 79 97
 DR DISULFID 91 108
 DR DISULFID 113 127
 DR DISULFID 120 140
 DR DISULFID 134 149
 DR DISULFID 154 166
 DR DISULFID 161 179
 DR DISULFID 173 188
 DR DISULFID 193 205
 DR DISULFID 200 218
 DR DISULFID 212 229
 DR DISULFID 239 251
 DR DISULFID 251 264
 DR DISULFID 268 273
 DR DISULFID 278 290
 DR DISULFID 285 303
 DR DISULFID 297 312
 DR DISULFID 318 331
 DR DISULFID 326 344
 DR DISULFID 338 355
 DR DISULFID 360 371
 DR DOMAIN 832 837
 DR DISULFID 33 45
 DR DISULFID 40 58
 DR DISULFID 52 67
 DR DISULFID 72 84
 DR DISULFID 79 97
 DR DISULFID 91 108
 DR DISULFID 113 127
 DR DISULFID 120 140
 DR DISULFID 134 149
 DR DISULFID 154 166
 DR DISULFID 161 179
 DR DISULFID 173 188
 DR DISULFID 193 205
 DR DISULFID 200 218
 DR DISULFID 212 229
 DR DISULFID 239 251
 DR DISULFID 251 264
 DR DISULFID 268 273
 DR DISULFID 278 290
 DR DISULFID 285 303
 DR DISULFID 297 312
 DR DISULFID 318 331
 DR DISULFID 326 344
 DR DISULFID 338 355
 DR DISULFID 360 371
 DR DOMAIN 832 837
 DR DISULFID 33 45
 DR DISULFID 40 58
 DR DISULFID 52 67
 DR DISULFID 72 84
 DR DISULFID 79 97
 DR DISULFID 91 108
 DR DISULFID 113 127
 DR DISULFID 120 140
 DR DISULFID 134 149
 DR DISULFID 154 166
 DR DISULFID 161 179
 DR DISULFID 173 188
 DR DISULFID 193 205
 DR DISULFID 200 218
 DR DISULFID 212 229
 DR DISULFID 239 251
 DR DISULFID 251 264
 DR DISULFID 268 273
 DR DISULFID 278 290
 DR DISULFID 285 303
 DR DISULFID 297 312
 DR DISULFID 318 331
 DR DISULFID 326 344
 DR DISULFID 338 355
 DR DISULFID 360 371
 DR DOMAIN 832 837
 DR DISULFID 33 45
 DR DISULFID 40 58
 DR DISULFID 52 67
 DR DISULFID 72 84
 DR DISULFID 79 97
 DR DISULFID 91 108
 DR DISULFID 113 127
 DR DISULFID 120 140
 DR DISULFID 134 149
 DR DISULFID 154 166
 DR DISULFID 161 179
 DR DISULFID 173 188
 DR DISULFID 193 205
 DR DISULFID 200 218
 DR DISULFID 212 229
 DR DISULFID 239 251
 DR DISULFID 251 264
 DR DISULFID 268 273
 DR DISULFID 278 290
 DR DISULFID 285 303
 DR DISULFID 297 312
 DR DISULFID 318 331
 DR DISULFID 326 344
 DR DISULFID 338 355
 DR DISULFID 360 371
 DR DOMAIN 832 837
 DR DISULFID 33 45
 DR DISULFID 40 58
 DR DISULFID 52 67
 DR DISULFID 72 84
 DR DISULFID 79 97
 DR DISULFID 91 108
 DR

FT	D1S0LFPD	120	140	BY SIMILARITY
FT	D1S0LFPD	134	149	BY SIMILARITY
FT	D1S0LFPD	154	166	BY SIMILARITY
FT	D1S0LFPD	161	179	BY SIMILARITY
FT	D1S0LFPD	173	188	BY SIMILARITY
FT	D1S0LFPD	193	205	BY SIMILARITY
FT	D1S0LFPD	200	218	BY SIMILARITY
FT	D1S0LFPD	212	229	BY SIMILARITY
FT	D1S0LFPD	239	251	BY SIMILARITY
FT	D1S0LFPD	246	264	BY SIMILARITY
FT	D1S0LFPD	258	273	BY SIMILARITY
FT	D1S0LFPD	278	290	BY SIMILARITY
FT	D1S0LFPD	285	303	BY SIMILARITY
FT	D1S0LFPD	297	312	BY SIMILARITY
FT	D1S0LFPD	318	331	BY SIMILARITY
FT	D1S0LFPD	326	344	BY SIMILARITY
FT	D1S0LFPD	338	355	BY SIMILARITY
FT	D1S0LFPD	360	371	BY SIMILARITY
FT	D1S0LFPD	367	380	BY SIMILARITY
FT	D1S0LFPD	382	394	BY SIMILARITY
FT	D1S0LFPD	400	419	BY SIMILARITY
FT	D1S0LFPD	406	419	BY SIMILARITY
FT	D1S0LFPD	421	434	BY SIMILARITY
FT	D1S0LFPD	706	719	BY SIMILARITY
FT	D1S0LFPD	715	734	BY SIMILARITY
FT	D1S0LFPD	736	749	BY SIMILARITY
FT	CARBOHYD	151	151	N-LINKED (GLCNAC . . .) (POTENTIAL)
FT	CARBOHYD	765	765	N-LINKED (GLCNAC . . .) (POTENTIAL)
FT	CARBOHYD	781	781	N-LINKED (GLCNAC . . .) (POTENTIAL)
FT	CONFLICT	161	161	C -> G (IN REF. 1)
FT	CONFLICT	263	263	P -> L (IN REF. 2)
FT	CONFLICT	297	297	C -> S (IN REF. 2)
NO	SEQUENCE	873 AA	96372 MW	9809P938251950ch CRC64

Query Match	17.08%	Score 82.5	DB 1	Length 873			
Host Local Similarity	41.78%	Pred. 0.59	Matches 26	Conservative 7	Mismatches 38	Indels 11	Gaps 4

UY	5	GCCKNEKYDS-----GSKKFC-DKCKKYGVVEFFNFFNPNVCTVWVHGLVWVFP	54
DB	228	GCWQVPIHKKPCTSPIDGDSPTGTHKKWPGGDDPKCKVSDSEVNVNCPSPCPQCECD	287
UY	55	GPYKNDKRCVACWAEFLNMD	76
DB	288	GCCLHNSGCTNGTTRFCVWDSID	308

RESULT 12			
LD LAYER PART	STANDARD:	PRT:	873 AA.
AC	P35953:		
DB	01-JUN-1994 (Ref. 29, Created)		
DB	01-JUN-1994 (Ref. 29, Last sequence update)		
DB	01-OCT-2000 (Ref. 40, Last annotation update)		
DB	VERY LOW-DENSITY LIPOPROTEIN RECEPTOR PRODUCIBLE (VLDL RECEPTOR)		
DB	VLDLR		
DB	Oryctolagus cuniculus (Rabbit).		
DB	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
DB	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.		
DB	Mol_taxid:9986;		
DB	111		
DB	SEQUENCE FROM N.A.		
DB	MEDLINE:93028432; PUBMED:184047.		
DB	Fukushashi S., Kawarabayashi Y., Nakai T., Sakai T., Yamamoto T.;		
DB	"Rabbit very low density lipoprotein receptor: a low density		
DB	lipoprotein receptor-like protein with distinct ligand specificity.";		
DB	Proc. Natl. Acad. Sci. U.S.A. 89:925-929(1992).		
DB	-1- FUNCTION: BINDS LDL AND TRANSPORTS IT INTO CELLS BY ENDOCYTOSIS.		
DB	IN ORDER TO BE INTERNALIZED, THE RECEPTOR-LIGAND COMPLEXES MUST		
DB	FIRST CLUSTER INTO CLATHRIN-COATED PITS		
DB	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.		
DB	-1- TISSUE SPECIFICITY: ABUNDANT IN HEART, MUSCLE, AND ADIPOSE		

CC	Tissue:
CC	-1- SIMILARITY: CONTAINS 8 LDL-RECEPTOR CLASS A DOMAINS.
CC	-1- SIMILARITY: CONTAINS 6 LDL-RECEPTOR CLASS B DOMAINS.
CC	-1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
CC	-----
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL collaboration with the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/aboutus/cv or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL: D11100; -1 NOT_ANNOTATED_CDS.
DR	PIR: A46286; ORFBN.
DR	HSSP: P01130; IAUJ.
DR	IleRfPro: IFRO00033; -1
DR	InterPro: IPRO000152; -1
DR	InterPro: IPRO000561; -1
DR	InterPro: IPRO01881; -1
DR	InterPro: IPRO02172; -1
DR	Pfam: PF000058; Ldl_recept_a: 8.
DR	Pfam: PF000058; Ldl_recept_b: 5.
DP	PRINTS: PR00261; LDLRECEPTOR.
DP	PROSITE: PS000910; ASX_HYDPOXYL_2.
DP	PROSITE: PS00022; EGF_1; FALSE_NEG.
DP	PROSITE: PS01146; EGF_2; 3.
DP	PROSITE: PS01147; EGF_CA_1.
DP	PROSITE: PS00009; LDLPAL_1; 8.
DP	PROSITE: PS00058; LDLPAL_2; 8.
KM	Glycoprotein VLDL cholesterol metabolism; lipid transport;
KM	Endocytosis; Coated pits; Transmembrane Receptor; Signal;
KM	EGF-like domain; Peptid-
FT	SIGNAL 1 27 POTENTIAL.
FT	CHAIN 28 873 VERY LOW-DENSITY LIPOPROTEIN RECEPTOR.
FT	DOMAIN 28 797 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 798 819 POTENTIAL.
FT	DOMAIN 820 873 CYTOSOLIC (POTENTIAL).
FT	DOMAIN 31 69 LDL-RECEPTOR CLASS A 1.
FT	DOMAIN 70 119 LDL-RECEPTOR CLASS A 2.
FT	DOMAIN 111 151 LDL-RECEPTOR CLASS A 3.
FT	DOMAIN 152 190 LDL-RECEPTOR CLASS A 4.
FT	DOMAIN 191 231 LDL-RECEPTOR CLASS A 5.
FT	DOMAIN 237 275 LDL-RECEPTOR CLASS A 6.
FT	DOMAIN 276 314 LDL-RECEPTOR CLASS A 7.
FT	DOMAIN 316 355 LDL-RECEPTOR CLASS A 8.
FT	DOMAIN 356 395 EGF-LIKE 1.
FT	DOMAIN 396 435 EGF-LIKE 2. CALCIUM-BINDING (POTENTIAL).
FT	REPEAT 439 480 LDL-RECEPTOR CLASS B 1.
FT	REPEAT 481 524 LDL-RECEPTOR CLASS B 2.
FT	REPEAT 525 567 LDL-RECEPTOR CLASS B 3.
FT	REPEAT 568 611 LDL-RECEPTOR CLASS B 4.
FT	REPEAT 612 654 LDL-RECEPTOR CLASS B 5.
FT	REPEAT 655 696 LDL-RECEPTOR CLASS B 6.
FT	DOMAIN 702 750 EGF-LIKE 3.
FT	DOMAIN 751 790 CLOUSTED Q-LINKED GLYCOCARPUS.
FT	DOMAIN 832 837 ENDOCYTOSIS SIGNAL (POTENTIAL).
FT	DISULFID 33 45 BY SIMILARITY.
FT	DISULFID 40 58 BY SIMILARITY.
FT	DISULFID 52 67 BY SIMILARITY.
FT	DISULFID 72 84 BY SIMILARITY.
FT	DISULFID 79 97 BY SIMILARITY.
FT	DISULFID 91 108 BY SIMILARITY.
FT	DISULFID 113 127 BY SIMILARITY.
FT	DISULFID 120 140 BY SIMILARITY.
FT	DISULFID 134 149 BY SIMILARITY.
FT	DISULFID 154 166 BY SIMILARITY.
FT	DISULFID 161 179 BY SIMILARITY.
FT	DISULFID 173 188 BY SIMILARITY.
FT	DISULFID 193 205 BY SIMILARITY.
FT	DISULFID 200 218 BY SIMILARITY.
FT	DISULFID 214 259 BY SIMILARITY.
FT	DISULFID 239 251 BY SIMILARITY.

FT	DISOLEP1D	246	264	BY SIMILARITY.
FT <th>DISOLEP1D</th> <th>258</th> <th>273</th> <th>BY SIMILARITY.</th>	DISOLEP1D	258	273	BY SIMILARITY.
FT <th>DISOLEP1D</th> <th>278</th> <th>290</th> <th>BY SIMILARITY.</th>	DISOLEP1D	278	290	BY SIMILARITY.
FT <th>DISOLEP1D</th> <th>285</th> <th>303</th> <th>BY SIMILARITY.</th>	DISOLEP1D	285	303	BY SIMILARITY.
FT <th>DISOLEP1D</th> <th>297</th> <th>312</th> <th>BY SIMILARITY.</th>	DISOLEP1D	297	312	BY SIMILARITY.
FT <th>DISOLEP1D</th> <th>318</th> <th>331</th> <th>BY SIMILARITY.</th>	DISOLEP1D	318	331	BY SIMILARITY.
FT <th>DISOLEP1D</th> <th>326</th> <th>344</th> <th>BY SIMILARITY.</th>	DISOLEP1D	326	344	BY SIMILARITY.
FT <th>DISOLEP1D</th> <th>338</th> <th>355</th> <th>BY SIMILARITY.</th>	DISOLEP1D	338	355	BY SIMILARITY.
FT <th>DISOLEP1D</th> <th>360</th> <th>371</th> <th>BY SIMILARITY.</th>	DISOLEP1D	360	371	BY SIMILARITY.
FT <th>DISOLEP1D</th> <th>367</th> <th>380</th> <th>BY SIMILARITY.</th>	DISOLEP1D	367	380	BY SIMILARITY.
FT <th>DISOLEP1D</th> <th>382</th> <th>394</th> <th>BY SIMILARITY.</th>	DISOLEP1D	382	394	BY SIMILARITY.
FT <th>DISOLEP1D</th> <th>400</th> <th>410</th> <th>BY SIMILARITY.</th>	DISOLEP1D	400	410	BY SIMILARITY.
FT <th>DISOLEP1D</th> <th>406</th> <th>419</th> <th>BY SIMILARITY.</th>	DISOLEP1D	406	419	BY SIMILARITY.
FT <th>DISOLEP1D</th> <th>421</th> <th>434</th> <th>BY SIMILARITY.</th>	DISOLEP1D	421	434	BY SIMILARITY.
FT <th>DISOLEP1D</th> <th>706</th> <th>719</th> <th>BY SIMILARITY.</th>	DISOLEP1D	706	719	BY SIMILARITY.
FT <th>DISOLEP1D</th> <th>715</th> <th>734</th> <th>BY SIMILARITY.</th>	DISOLEP1D	715	734	BY SIMILARITY.
FT <th>DISOLEP1D</th> <th>736</th> <th>749</th> <th>BY SIMILARITY.</th>	DISOLEP1D	736	749	BY SIMILARITY.
FT <th>CARBONYD</th> <th>151</th> <th>151</th> <th>N-LINKED (GLUCNAC. . .) (POTENTIAL).</th>	CARBONYD	151	151	N-LINKED (GLUCNAC. . .) (POTENTIAL).
FT <th>CARBONYD</th> <th>765</th> <th>765</th> <th>N-LINKED (GLUCNAC. . .) (POTENTIAL).</th>	CARBONYD	765	765	N-LINKED (GLUCNAC. . .) (POTENTIAL).
FT <th>CARBONYD</th> <th>781</th> <th>781</th> <th>N-LINKED (GLUCNAC. . .) (POTENTIAL).</th>	CARBONYD	781	781	N-LINKED (GLUCNAC. . .) (POTENTIAL).
FT <th>SEQUENCE</th> <th>873 AA.</th> <th>946-893 MW.</th> <th>CP55441.251A1933 CP664.</th>	SEQUENCE	873 AA.	946-893 MW.	CP55441.251A1933 CP664.

```

Query Match: 17.0%; Score 82.5; DB 1; Length 873;
Best Local Similarity: 41.7%; Prod No 0 99;
Matches 25; Conservative 7; Mismatches 38; Indels 11; Gaps 4.

QY 5 GCSENEKYVS-----CGSKEC-DKKCKYDVGVEEDEPPWCVIVKCHOD-CVCEE 54
LD 228 GCPGCVVTHFETVAST-CTGTSYVIRKSWP-IVGATGKSLNLEAVYVTSFVETGIVPPTD 287
DB 298 GCSLHSPGNCJIPG-VPSPD 308

QY 55 GEFYRKDKRQVSAEPELJLNNQ 76
DB 298 GCSLHSPGNCJIPG-VPSPD 308

RESULT 13
LDVR_RAT
ID LDVR_RAT STANDARD: PRT: 873 AA.
AC P88166;
DI 01-001-1996 (Rel. 34, Created)
DI 01-001-1996 (Rel. 34, last sequence update)
DI 01-001-2000 (Rel. 40, last annotation update)
DE VERY LOW-DENSITY LIPOPROTEIN RECEPTOR PRECURSOR (VLDL RECEPTOR).
GN VLDLR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia, Sciurognathi, Muridae, Murinae, Felinae.
OX NCBI_Taxid:10116;
KP SPIDENCE FROM N.A.
RC 11SSDE-Brain;
RX MED1 (NP_05014480); P01M04-792962;
RA Johnson E V., Landseboy F T., Wynn F L., Ho Y K., Frykner P.K.,
RA Hobbs H.H.;
RT "Regulation of the very low density lipoprotein receptor by thyroid
RT hormone in rat skeletal muscle."
PI J Biol Chem 266:26411-26418(1994)
CC -1- FUNCTION: BINDS VLDL AND TRANSPORTS IT INTO CELLS BY ENDOCYTOSIS.
CC IN ORDER TO BE INTERNALIZED, THE RECEPTOR-LIGAND COMPLEXES MUST
CC FIRST CLUSTER AND INTERNALIZE INTO CLATHRIN-COATED PITS.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: ABUNDANT IN SOLEUS, GASTROCNEMIUS, HEART
CC MUSCLE, PLACENTA, BRAIN, LUNG AND WHITE FAT. LESS IN ADIPI, OVARY,
CC KIDNEY, SPLEEN, ADRENAL GLAND, AND THYMUS.
CC -1- SIMILARITY: CONTAINS 8 LDL-RECEPTOR CLASS A DOMAINS
CC -1- SIMILARITY: CONTAINS 6 LDL-RECEPTOR CLASS B DOMAINS.
CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
CC
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL collaboration.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```


CC entities requires a license agreement (See <http://www.isb-sib.ch/announcements/>).

CC or send an email to license@isb-sib.ch).

EMBL: M23121: AAA4086.1: -

EMBL: X08031: CAB4081.1: -

EMBL: X08030: CAB3082.1: ALT_TERM.

EMBL: J07641: AAA4874.1: ALT_SEQ.

EMBL: M20816: AAA4874.1: ALT_SEQ.

PIR: A30903: A30903.

PIR: A31930: A31930.

PIR: A33379: A33379.

PIR: B33379: B33379.

PIR: C33379: C33379.

PIR: S01292: S01292.

HSSP: P24821: ITEN.

InterPro: IP000561: -

InterPro: IP001777: -

InterPro: IP002181: -

PIR: P00008: EGF_13.

PIR: P00147: Fibrinogen_C_1.

PIR: P00041: fn3_11.

PROSITE: PS00022: EGF_1: 14.

PROSITE: PS01186: EGF_2: 14.

GLycoprotein: Cell adhesion; Repeat; EGF-like domain; Coiled coil; Extracellular matrix; Signal; Alternative splicing.

KW Extracellular matrix; Signal; Alternative splicing.

FT SIGNAL 1 22

FT PROPEP 23 33

FT CHAIN 34 1808

FT DOMAIN 118 142

FT DOMAIN 176 188

FT DOMAIN 188 219

FT DOMAIN 219 250

FT DOMAIN 250 281

FT DOMAIN 281 312

FT DOMAIN 312 343

FT DOMAIN 343 374

FT DOMAIN 374 405

FT DOMAIN 405 436

FT DOMAIN 436 467

FT DOMAIN 467 498

FT DOMAIN 498 529

FT DOMAIN 529 560

FT DOMAIN 560 591

FT DOMAIN 591 621

FT DOMAIN 621 652

FT DOMAIN 652 683

FT DOMAIN 683 714

FT DOMAIN 714 745

FT DOMAIN 745 776

FT DOMAIN 776 807

FT DOMAIN 807 838

FT DOMAIN 838 869

FT DOMAIN 869 900

FT DOMAIN 900 931

FT DOMAIN 931 962

FT DOMAIN 962 993

FT DOMAIN 993 1024

FT DOMAIN 1024 1055

FT DOMAIN 1055 1086

FT DOMAIN 1086 1117

FT DOMAIN 1117 1148

FT DOMAIN 1148 1179

FT DOMAIN 1179 1210

FT DOMAIN 1210 1241

FT DOMAIN 1241 1272

FT DOMAIN 1272 1303

FT DOMAIN 1303 1334

FT DOMAIN 1334 1365

FT DOMAIN 1365 1396

FT DOMAIN 1396 1427

FT DOMAIN 1427 1458

FT DOMAIN 1458 1489

FT DOMAIN 1489 1520

FT DOMAIN 1520 1551

FT DOMAIN 1551 1582

FT DOMAIN 1582 1613

FT DOMAIN 1613 1644

FT DOMAIN 1644 1675

FT DOMAIN 1675 1706

FT DOMAIN 1706 1737

FT DOMAIN 1737 1768

FT DOMAIN 1768 1799

FT DOMAIN 1799 1830

FT DOMAIN 1830 1861

FT DOMAIN 1861 1892

FT DOMAIN 1892 1923

FT DOMAIN 1923 1954

FT DOMAIN 1954 1985

FT DOMAIN 1985 2016

FT DOMAIN 2016 2047

FT DOMAIN 2047 2078

FT DOMAIN 2078 2109

FT DOMAIN 2109 2140

FT DOMAIN 2140 2171

FT DOMAIN 2171 2202

FT DOMAIN 2202 2233

FT DOMAIN 2233 2264

FT DOMAIN 2264 2295

FT DOMAIN 2295 2326

FT DOMAIN 2326 2357

FT DOMAIN 2357 2388

FT DOMAIN 2388 2419

FT DOMAIN 2419 2450

FT DOMAIN 2450 2481

FT DOMAIN 2481 2512

FT DOMAIN 2512 2543

FT DOMAIN 2543 2574

FT DOMAIN 2574 2605

FT DOMAIN 2605 2636

FT DOMAIN 2636 2667

FT DOMAIN 2667 2698

FT DOMAIN 2698 2729

FT DOMAIN 2729 2760

FT DOMAIN 2760 2791

FT DOMAIN 2791 2822

FT DOMAIN 2822 2853

FT DOMAIN 2853 2884

FT DOMAIN 2884 2915

FT DOMAIN 2915 2946

FT DOMAIN 2946 2977

FT DOMAIN 2977 3008

FT DOMAIN 3008 3039

FT DOMAIN 3039 3070

FT DOMAIN 3070 3101

FT DOMAIN 3101 3132

FT DOMAIN 3132 3163

FT DOMAIN 3163 3194

FT DOMAIN 3194 3225

FT DOMAIN 3225 3256

FT DOMAIN 3256 3287

FT DOMAIN 3287 3318

FT DOMAIN 3318 3349

FT DOMAIN 3349 3380

FT DOMAIN 3380 3411

FT DOMAIN 3411 3442

FT DOMAIN 3442 3473

FT DOMAIN 3473 3504

FT DOMAIN 3504 3535

FT DOMAIN 3535 3566

FT DOMAIN 3566 3597

FT DOMAIN 3597 3628

FT DOMAIN 3628 3659

FT DOMAIN 3659 3690

FT DOMAIN 3690 3721

FT DOMAIN 3721 3752

FT DOMAIN 3752 3783

FT DOMAIN 3783 3814

FT DOMAIN 3814 3845

FT DOMAIN 3845 3876

FT DOMAIN 3876 3907

FT DOMAIN 3907 3938

FT DOMAIN 3938 3969

FT DOMAIN 3969 4000

FT DOMAIN 4000 4031

FT DOMAIN 4031 4062

FT DOMAIN 4062 4093

FT DOMAIN 4093 4124

FT DOMAIN 4124 4155

FT DOMAIN 4155 4186

FT DOMAIN 4186 4217

FT DOMAIN 4217 4248

FT DOMAIN 4248 4279

FT DOMAIN 4279 4310

FT DOMAIN 4310 4341

FT DOMAIN 4341 4372

FT DOMAIN 4372 4403

FT DOMAIN 4403 4434

FT DOMAIN 4434 4465

FT DOMAIN 4465 4496

FT DOMAIN 4496 4527

FT DOMAIN 4527 4558

FT DOMAIN 4558 4589

FT DOMAIN 4589 4620

FT DOMAIN 4620 4651

FT DOMAIN 4651 4682

FT DOMAIN 4682 4713

FT DOMAIN 4713 4744

FT DOMAIN 4744 4775

FT DOMAIN 4775 4806

FT DOMAIN 4806 4837

FT DOMAIN 4837 4868

FT DOMAIN 4868 4899

FT DOMAIN 4899 4930

FT DOMAIN 4930 4961

FT DOMAIN 4961 4992

FT DOMAIN 4992 5023

FT DOMAIN 5023 5054

FT DOMAIN 5054 5085

FT DOMAIN 5085 5116

FT DOMAIN 5116 5147

FT DOMAIN 5147 5178

FT DOMAIN 5178 5209

FT DOMAIN 5209 5240

FT DOMAIN 5240 5271

FT DOMAIN 5271 5302

FT DOMAIN 5302 5333

FT DOMAIN 5333 5364

FT DOMAIN 5364 5395

FT DOMAIN 5395 5426

FT DOMAIN 5426 5457

FT DOMAIN 5457 5488

FT DOMAIN 5488 5519

FT DOMAIN 5519 5550

FT DOMAIN 5550 5581

FT DOMAIN 5581 5612

FT DOMAIN 5612 5643

FT DOMAIN 5643 5674

FT DOMAIN 5674 5705

FT DOMAIN 5705 5736

FT DOMAIN 5736 5767

FT DOMAIN 5767 5798

FT DOMAIN 5798 5829

FT DOMAIN 5829 5860

FT DOMAIN 5860 5891

FT DOMAIN 5891 5922

FT DOMAIN 5922 5953

FT DOMAIN 5953 5984

FT DOMAIN 5984 6015

FT DOMAIN 6015 6046

FT DOMAIN 6046 6077

FT DOMAIN 6077 6108

FT DOMAIN 6108 6139

FT DOMAIN 6139 6170

FT DOMAIN 6170 6201

FT DOMAIN 6201 6232

FT DOMAIN 6232 6263

FT DOMAIN 6263 6294

FT DOMAIN 6294 6325

FT DOMAIN 6325 6356

FT DOMAIN 6356 6387

FT DOMAIN 6387 6418

FT DOMAIN 6418 6449

FT DOMAIN 6449 6480

FT DOMAIN 6480 6511

FT DOMAIN 6511 6542

FT DOMAIN 6542 6573

FT DOMAIN 6573 6604

FT DOMAIN 6604 6635

FT DOMAIN 6635 6666

FT DOMAIN 6666 6697

FT DOMAIN 6697 6728

FT DOMAIN 6728 6759

FT DOMAIN 6759 6790

FT DOMAIN 6790 6821

FT DOMAIN 6821 6852

FT DOMAIN 6852 6883

FT DOMAIN 6883 6914

FT DOMAIN 6914 6945

FT DOMAIN 6945 6976

FT DOMAIN 6976 7007

FT DOMAIN 7007 7038

FT DOMAIN 7038 7069

FT DOMAIN 7069 7100

FT DOMAIN 7100 7131

FT DOMAIN 7131 7162

FT DOMAIN 7162 7193

FT DOMAIN 7193 7224

FT DOMAIN 7224 7255

FT DOMAIN 7255 7286

FT DOMAIN 7286 7317

FT DOMAIN 7317 7348

FT DOMAIN 7348 7379

FT DOMAIN 7379 7410

FT DOMAIN 7410 7441

FT DOMAIN 7441 7472

FT DOMAIN 7472 7503

FT DOMAIN 7503 7534

FT DOMAIN 7534 7565

FT DOMAIN 7565 7596

FT DOMAIN 7596 7627

FT DOMAIN 7627 7658

FT DOMAIN 7658 7689

FT DOMAIN 7689 7720

FT DOMAIN 7720 7751

FT DOMAIN 7751 7782

FT DOMAIN 7782 7813

FT DOMAIN 7813 7844

FT DOMAIN 7844 7875

FT DOMAIN 7875 7906

FT DOMAIN 7906 7937

FT DOMAIN 7937 7968

FT DOMAIN 7968 7999

FT DOMAIN 7999 8030

FT DOMAIN 8030 8061

FT DOMAIN 8061 8092

FT DOMAIN 8092 8123

FT DOMAIN 8123 8154

FT DOMAIN 8154 8185

FT DOMAIN 8185 8216

FT DOMAIN 8216 8247

FT DOMAIN 8247 8278

FT DOMAIN 8278 8309

FT DOMAIN 8309 8340

FT DOMAIN 8340 8371

FT DOMAIN 8371 8402

FT DOMAIN 8402 8433

FT DOMAIN 8433 8464

FT DOMAIN 8464 8495

FT DOMAIN 8495 8526

FT DOMAIN 8526 8557

FT DOMAIN 8557 8588

FT DOMAIN 8588 8619

FT DOMAIN 8619 8650

FT DOMAIN 8650 8681

FT DOMAIN 8681 8712

FT DOMAIN 8712 8743

FT DOMAIN 8743 8774

FT DOMAIN 8774 8805

FT DOMAIN 8805 8836

FT DOMAIN 8836 8867

FT DOMAIN 8867 8898

FT DOMAIN 8898 8929

FT DOMAIN 8929 8960

FT DOMAIN 8960 8991

FT DOMAIN 8991 9022

FT DOMAIN 9022 9053

FT DOMAIN 9053 9084

FT DOMAIN 9084 9115

FT DOMAIN 9115 9146

FT DOMAIN 9146 9177

FT DOMAIN 9177 9208

FT DOMAIN 9208 9239

FT DOMAIN 9239 9270

FT DOMAIN 9270 9301

FT DOMAIN 9301 9332

FT DOMAIN 9332 9363

FT DOMAIN 9363 9394

FT DOMAIN 9394 9425

FT DOMAIN 9425 9456

FT DOMAIN 9456 9487

FT DOMAIN 9487 9518

FT DOMAIN 9518 9549

FT DOMAIN 9549 9580

FT DOMAIN 9580 9611

FT DOMAIN 9611 9642

FT DOMAIN 9642 9673

FT DOMAIN 9673 9704

FT DOMAIN 9704 9735

FT DOMAIN 9735 9766

FT DOMAIN 9766 9797

FT DOMAIN 9797 9828

FT DOMAIN 9828 9859

FT DOMAIN 9859 9890

FT DOMAIN 9890 9921

FT DOMAIN 9921 9952

FT DOMAIN 9952 9983

FT DOMAIN 9983 10014

FT DOMAIN 10014 10045

FT DOMAIN 10045 10076

FT DOMAIN 10076 10107

FT DOMAIN 10107 10138

FT DOMAIN 10138 10169

FT DOMAIN 10169 10200

FT DOMAIN 10200 10231

FT DOMAIN 10231 10262

FT DOMAIN 10262 10293

FT DOMAIN 10293 10324

FT DOMAIN 10324 10355

FT DOMAIN 10355 10386

FT DOMAIN 10386 10417

FT DOMAIN 10417 10448

FT DOMAIN 10448 10479

FT DOMAIN 10479 10510

FT DOMAIN 10510 10541

FT DOMAIN 10541 10572

FT DOMAIN 10572 10603

FT DOMAIN 10603 10634

FT DOMAIN 10634 10665

FT DOMAIN 10665 10696

FT DOMAIN 10696 10727

FT DOMAIN 10727 10758

FT DOMAIN 10758 10789

FT DOMAIN 10789 10820

FT DOMAIN 10820 10851

FT DOMAIN 10851 10882

FT DOMAIN 10882 10913

FT DOMAIN 10913 10944

FT DOMAIN 10944 10975

FT DOMAIN 10975 11006

FT DOMAIN 11006 11037

FT DOMAIN 11037 11068

FT DOMAIN 11068 11099

FT DOMAIN 11099 11130

FT DOMAIN 11130 11161

FT DOMAIN 11161 11192

FT DOMAIN 11192 11223

FT DOMAIN 11223 11254

FT DOMAIN 11254 11285

FT DOMAIN 11285 11316

FT DOMAIN 11316 11347

FT DOMAIN 11347 11378

FT DOMAIN 11378 11409

FT DOMAIN 11409 11440

FT DOMAIN 11440 11471

FT DOMAIN 11471 11502

FT DOMAIN 11502 11533

FT DOMAIN 11533 11564

FT DOMAIN 11564 11595

FT DOMAIN 11595 11626

FT DOMAIN 11626 11657

FT DOMAIN 11657 11688

FT DOMAIN 11688 11719

FT DOMAIN 11719 11750

FT DOMAIN 11750 11781

FT DOMAIN 11781 11812

FT DOMAIN 11812 11843

FT DOMAIN 11843 11874

FT DOMAIN 11874 11905

FT DOMAIN 11905 11936

FT DOMAIN 11936 11967

FT DOMAIN 11967 11998

FT DOMAIN 11998 12029

FT DOMAIN 12029 12060

FT DOMAIN 12060 12091

FT DOMAIN 12091 12122

FT DOMAIN 12122 12153

FT DOMAIN 12153 12184

FT DOMAIN 12184 12215

FT DOMAIN 12215 12246

FT DOMAIN 12246 12277

FT DOMAIN 12277 12308

FT DOMAIN 12308 12339

FT DOMAIN 12339 12370

FT DOMAIN 12370 12401

FT DOMAIN 12401 12432

FT DOMAIN 12432 12463

FT DOMAIN 12463 12494

FT DOMAIN 12494 12525

FT DOMAIN 12525 12556

FT DOMAIN 12556 12587

FT DOMAIN 12587 12618

FT DOMAIN 12618 12649

FT DOMAIN 12649 12680

FT DOMAIN 12680 12711

FT DOMAIN 12711 12742

FT DOMAIN 12742 12773

FT DOMAIN 12773 12804

FT DOMAIN 12804 12835

FT DOMAIN 12835 12866

FT DOMAIN 12866 12897

FT DOMAIN 12897 12928

FT DOMAIN 12928 12959

FT DOMAIN 12959 12990

FT DOMAIN 12990 13021

FT DOMAIN 13021 13052

FT DOMAIN 13052 13083

FT DOMAIN 13083 13114

FT DOMAIN 13114 13145

FT DOMAIN 13145 13176

FT DOMAIN 13176 13207

FT DOMAIN 13207 13238

FT DOMAIN 13238 13269

FT DOMAIN 13269 13300

FT DOMAIN 13300 13331

FT DOMAIN 13331 13362

FT DOMAIN 13362 13393

FT DOMAIN 13393 13424

FT DOMAIN 13424 13455

FT DOMAIN 13455 13486

FT DOMAIN 13486 13517

FT DOMAIN 13517 13548

FT DOMAIN 13548 13579

FT DOMAIN 13579 13610

FT DOMAIN 13610 13641

FT DOMAIN 13641 13672

FT DOMAIN 13672 13703

FT DOMAIN 13703 13734

FT DOMAIN 13734 13765

FT DOMAIN 13765 13796

FT DOMAIN 13796 13827

FT DOMAIN 13827 13858

FT DOMAIN 13858 13889

FT DOMAIN 13889 13920

FT DOMAIN 13920 13951

FT DOMAIN 13951 13982

FT DOMAIN 13982 14013

FT DOMAIN 14013 14044

FT DOMAIN 14044 14075

FT DOMAIN 14075 14106

FT DOMAIN 14106 14137

FT DOMAIN 14137 14168

FT DOMAIN 14168 14199

FT DOMAIN 14199 14230

FT DOMAIN 14230 14261

FT DOMAIN 14261 14292

FT DOMAIN 14292 14323

FT DOMAIN 14323 14354

FT DOMAIN 14354 14385

FT DOMAIN 14385 14416

FT DOMAIN 14416 14447

FT DOMAIN 14447 14478

FT DOMAIN 14478 14509

FT DOMAIN 14509 14540

FT DOMAIN 14540 14571

FT DOMAIN 14571 14602

FT DOMAIN 14602 14633

FT DOMAIN 14633 14664

FT DOMAIN 14664 14695

FT DOMAIN 14695 14726

FT DOMAIN 14726 14757

FT DOMAIN 14757 14788

FT DOMAIN 14788 14819

FT DOMAIN 14819 14850

FT DOMAIN 14850 14881

FT DOMAIN 14881 14912

FT DOMAIN 14912 14943

FT DOMAIN 14943 14974

FT DOMAIN 14974 15005

FT DOMAIN 15005 15036

FT DOMAIN 15036 15067

FT DOMAIN 15067 15098

FT DOMAIN 15098 15129

FT DOMAIN 15129 15160

FT DOMAIN 15160 15191

FT DOMAIN 15191 15222

FT DOMAIN 15222 15253

FT DOMAIN 15253 15284

FT DOMAIN 15284 15315

FT DOMAIN 15315 15346

FT DOMAIN 15346 15377

FT DOMAIN 15377 15408

FT DOMAIN 15408 15439

FT DOMAIN 15439 15470

FT DOMAIN 15470 15501

FT DOMAIN 15501 15532

FT DOMAIN 15532 15563

FT DOMAIN 15563 15594

FT DOMAIN 15594 15625

FT DOMAIN 15625 15656

FT DOMAIN 15656 15687

FT DOMAIN 15687 15718

FT DOMAIN 15718 15749

FT DOMAIN 15749 15780

FT DOMAIN 15780 15811

FT DOMAIN 15811 15842

FT DOMAIN 15842 15873

FT DOMAIN 15873 15904

FT DOMAIN 15904 15935

FT DOMAIN 15935 15966

FT DOMAIN 15966 15997

FT DOMAIN 15997 16028

FT DOMAIN 16028 16059

FT DOMAIN 16059 16090

FT DOMAIN 16090 16121

FT DOMAIN 16121 16152

FT DOMAIN 16152 16183

FT DOMAIN 16183 16214

FT DOMAIN 16214 16245

FT DOMAIN 16245 16276

FT DOMAIN 16276 16307

FT DOMAIN 16307 16338

FT DOMAIN 16338 16369

FT DOMAIN 16369 16400

FT DOMAIN 16400 16431

FT DOMAIN 16431 16462

FT DOMAIN 16462 16493

FT DOMAIN 16493 16524

FT DOMAIN 16524 16555

FT DOMAIN 16555 16586

FT DOMAIN 16586 16617

FT DOMAIN 16617 16648

FT DOMAIN 16648 16679

FT DOMAIN 16679 16710

FT DOMAIN 16710 16741

FT DOMAIN 16741 16772

FT DOMAIN 16772 16803

FT DOMAIN 16803 16834

FT DOMAIN 16834 16865

FT DOMAIN 16865 16896

FT DOMAIN 16896 16927

FT DOMAIN 16927 16958

FT DOMAIN 16958 16989

FT DOMAIN 16989 17020

FT DOMAIN 17020 17051

FT DOMAIN 17051 17082

FT DOMAIN 17082 17113

FT DOMAIN 17113 17144

FT DOMAIN 17144 17175

FT DOMAIN 17175 17206

FT DOMAIN 17206 17237

FT DOMAIN 17237 17268

FT DOMAIN 17268 17299

FT DOMAIN 17299 17330

FT DOMAIN 17330 17361

FT DOMAIN 17361 17392

FT DOMAIN 17392 17423

FT DOMAIN 17423 17454

FT DOMAIN 17454 17485

FT DOMAIN 17485 17516

FT DOMAIN 17516 17547

FT DOMAIN 17547 17578

FT DOMAIN 17578 17609

FT DOMAIN 17609 17640

FT DOMAIN 17640 17671

FT DOMAIN 17671 17702

FT DOMAIN 17702 17733

FT DOMAIN 17733 17764

FT DOMAIN 17764 17795

FT DOMAIN 17795 17826

FT DOMAIN 17826 17857

FT DOMAIN 17857 17888

FT DOMAIN 17888 17919

FT DOMAIN 17919 17950

FT DOMAIN 17950 17981

FT DOMAIN 17981 18012

FT DOMAIN 18012 18043

FT DOMAIN 18043 18074

FT DOMAIN 18074 18105

FT DOMAIN 18105 18136

FT DOMAIN 18136 18167

FT DOMAIN 18167 18198

FT DOMAIN 18198 18229

FT DOMAIN 18229 18260

FT DOMAIN 18260 18291

FT DOMAIN 18291 18322

FT DOMAIN 18322 18353

FT DOMAIN 18353 18384

FT DOMAIN 18384 18415

FT DOMAIN 18415 18446

FT DOMAIN 18446 18477

FT DOMAIN 18477 18508

FT DOMAIN 18508 18539

FT DOMAIN 18539 18570

FT DOMAIN 18570 18601

FT DOMAIN 18601 18632

FT DOMAIN 18632 18663

FT DOMAIN 18663 18694

FT DOMAIN 18694 18725

FT DOMAIN 18725 18756

FT DOMAIN 18756 18787

FT DOMAIN 18787 18818

FT DOMAIN 18818 18849

FT DOMAIN 18849 18880

FT DOMAIN 18880 18911

FT DOMAIN 18911 18942

FT DOMAIN 18942 18973

FT DOMAIN 18973 19004

FT DOMAIN 19004 19035

FT DOMAIN 19035 19066

FT DOMAIN 19066 19097

FT DOMAIN 19097 19128

FT DOMAIN 19128 19159

FT DOMAIN 19159 19190

US-08-486-399-59

Query Match: 100.0% Score 486; DB 2; Length 84;

Post Local Similarity 100.0% Prod. No. 2,26-43; Matches 84; Consolative 0; Mismatches 0; Indels 0; Gaps 0;

1 1 KATMGENEKEKESKEDJOKKVEVEEEDDEPWOTVWVTHQWVVEEVEENK 60
 1 KATMGENEKEKESKEDJOKKVEVEEEDDEPWOTVWVTHQWVVEEVEENK 60

61 DOKVSAHE*TELEMDIYVSHN 84
 61 DOKVSAHE*TELEMDIYVSHN 84

US-08-486-399-59
 Sequence 59; Application 05/08486/59

Patent No. 5866542
 GENERAL INFORMATION:

APPLICANT: George P. Vlasok, Patrick H. Stansons,

APPLICANT: Joris H.L. Meusens, Marc J. Lamberys,

APPLICANT: Yves R. Laroche, Laurent St. Josephs,

APPLICANT: Yvanick G.J. Gaussemans, Matthew Moyte,

APPLICANT: Peter W. Bequim

TITLE OF INVENTION: REMAINDER EXTRACTED ANTICOAGULANT

NUMBER OF SEQUENCES: 457

CORRESPONDENT ADDRESS:

ADDRESS: Lyon & Lyon

STREET: 644 West Fifth Street

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 4.5" Diskette, 1.44 MB

MEDIUM TYPE: Storage

COMPUTER: IBM compatible

OPERATING SYSTEM: IBM PC, DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: 08/026,110

FILING DATE: October 18, 1994

ATTORNEY/AGENT INFORMATION:

NAME: BROS. SOZANNE L.

REGISTRATION NUMBER: 40,158

REFERENCE: 5,871,226; 5,871,227

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-4910

INFORMATION FOR SEQ ID NO: 59:

SEQUENCE CHARACTERISTICS:

LENGTH: 84 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

ORIGINAL SOURCE:

ORGANISM: Ancylostoma caninum

US-08-486-399-59

Query Match: 100.0% Score 486; DB 2; Length 84;

Post Local Similarity 100.0% Prod. No. 2,26-43;

Matches 84; Consolative 0; Mismatches 0; Indels 0; Gaps 0;

Query 1 KATMGENEKEKESKEDJOKKVEVEEEDDEPWOTVWVTHQWVVEEVEENK 60

DB 1 KATMGENEKEKESKEDJOKKVEVEEEDDEPWOTVWVTHQWVVEEVEENK 60

Query 61 DOKVSAHE*TELEMDIYVSHN 84

DB 61 DOKVSAHE*TELEMDIYVSHN 84

RESULT 3

US-08-486-399-59

Sequence 59; Application 05/08486/59

Patent No. 5866542

GENERAL INFORMATION:

APPLICANT: George P. Vlasok, Patrick H. Stansons,

APPLICANT: Joris H.L. Meusens, Marc J. Lamberys,

APPLICANT: Yves R. Laroche, Laurent St. Josephs,

APPLICANT: Yvanick G.J. Gaussemans, Matthew Moyte,

APPLICANT: Peter W. Bequim

TITLE OF INVENTION: REMAINDER EXTRACTED ANTICOAGULANT

NUMBER OF SEQUENCES: 456

CORRESPONDENT ADDRESS:

ADDRESS: Lyon & Lyon

STREET: 644 West Fifth Street

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 4.5" Diskette, 1.44 MB

MEDIUM TYPE: Storage

COMPUTER: IBM compatible

OPERATING SYSTEM: IBM PC, DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: 08/026,110

FILING DATE: October 18, 1994

ATTORNEY/AGENT INFORMATION:

NAME: BROS. SOZANNE L.

REGISTRATION NUMBER: 40,158

REFERENCE: 5,871,226; 5,871,227

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-4910

INFORMATION FOR SEQ ID NO: 59:

SEQUENCE CHARACTERISTICS:

LENGTH: 84 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

ORIGINAL SOURCE:

ORGANISM: Ancylostoma caninum

US-08-486-399-59

Query Match: 100.0% Score 486; DB 2; Length 84;

Post Local Similarity 100.0% Prod. No. 2,26-43;

Matches 84; Consolative 0; Mismatches 0; Indels 0; Gaps 0;

Query 61 DOKVSAHE*TELEMDIYVSHN 84

DB 61 DOKVSAHE*TELEMDIYVSHN 84

1 PUBLISH DATE: JUN 5, 1995
 2 APPLICATION NUMBER: 09/426,110
 3 FILING DATE: OCTOBER 18, 1994
 4 ATTORNEY/AGENT INFORMATION:
 5 NAME: BLOIS, STEVEN L
 6 REGISTRATION NUMBER: 40,150
 7 KEYWORD(S): 500, 506, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 841, 842, 843, 844, 845, 846, 847, 848, 849, 850, 851, 852, 853, 854, 855, 856, 857, 858, 859, 860, 861, 862, 863, 864, 865, 866, 867, 868, 869, 870, 871, 872, 873, 874, 875, 876, 877, 878, 879, 880, 881, 882, 883, 884, 885, 886, 887, 888, 889, 890, 891, 892, 893, 894, 895, 896, 897, 898, 899, 900, 901, 902, 903, 904, 905, 906, 907, 908, 909, 910, 911, 912, 913, 914, 915, 916, 917, 918, 919, 920, 921, 922, 923, 924, 925, 926, 927, 928, 929, 930, 931, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, 972, 973, 974, 975, 976, 977, 978, 979, 980, 981, 982, 983, 984, 985, 986, 987, 988, 989, 990, 991, 992, 993, 994, 995, 996, 997, 998, 999, 1000
 9 ORGANISM: *Myxocystis xanthus*

```

1 COMPUTER: IBM Compatible
2 OPERATING SYSTEM: IBM 4.2 iOS 5.0
3 SOFTWARE: Word Perfect 5.1
4 CURRENT APPLICATION DATA:
5 AFFILIATION NUMBER: 00707219, 416
6 FILING DATE:
7 PAPER APPLICATION DATA:
8 APPLICATION NUMBER: 08/2009, 455
9 FILING DATE: April 15, 1995
10 AFFILIATION NUMBER: 01/0895, 72231
11 FILING DATE: October 17, 1995
12 AFFILIATION NUMBER: 08/155, 594
13 FILING DATE: June 5, 1995
14 APPLICATION NUMBER: 08/486, 997
15 FILING DATE: June 5, 1995
16 AFFILIATION NUMBER: 06/105, 480
17 FILING DATE: June 5, 1995
18 AFFILIATION NUMBER: 08/361, 965
19 FILING DATE: June 5, 1995
20 APPLICATION NUMBER: 08/326, 110
21 FILING DATE: October 18, 1994
22 ATTORNEY/AGENT INFORMATION:
23 NAME: BROWN, STEWARD L.
24 REGISTRATION NUMBER: 30,158
25 REFERENCE/2: FILL REQUEST: 215, 570
26 TELECOMMUNICATION INFORMATION:
27 TELEPHONE: (213) 489-1600
28 TELEFAX: (213) 955-0440
29 TELE: 67-8510
30 INFORMATION FOR SEQ ID NO. 59:
31 SEQUENCE CHARACTERISTICS:
32 LENGTH: 84 amino acids
33 TYPE: amino acid
34 TOPOLOGY: linear
35 MAPPING: linear
36 ORIGINAL SOURCE:
37 ORGANISM: Anacylostoma caninum
38 DB: 249-448 59
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
9
```

```

1 ZIP: 90071
2 COMPUTER READABLE FORM:
3 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
4 MEDIUM TYPE: storage
5 MEDIUM TYPE: COMPATIBLE
6 OPERATING SYSTEM: IBM PC, DOS 5.0
7 SOFTWARE: word perfect 5.1
8 CURRENT APPLICATION DATA:
9 APPLICATION NUMBER: 05/98/165,380
10 FILING LABEL: June 5, 1995
11 CLASSIFICATION: 530
12 FOR APPLICATION DATA:
13 APPLICATION NUMBER: 08/26,110
14 FILING DATE: October 18, 1994
15 AUTOMATIC/ANTI INFORMATION:
16 NAME: POLIO, JOHANN L.
17 REGISTRATION NUMBER: 40,158
18 FEE/FILE/CLASS: NONE, 13,268
19 TELECOMMUNICATION INFORMATION:
20 TELEPHONE: (213) 489,1600
21 TELEFAX: (213) 955-0440
22 TELEX: 47-3510
23 INFORMATION FOR SEQ ID NO: 128:
24 SEQUENCE CHARACTERISTICS:
25 LENGTH: 91 amino acids
26 TYPE: amino acid
27 TOPOLOGY: linear
28 MOLECULE TYPE: peptide
29 ORIGINAL SOURCE:
30 ORGANISM: Ancylostoma caninum
31 US-08-465-480-128
32
33 Query Match: 100.0%; Score 486; Dh 2; Length 91;
34 Best Local Similarity: 100.0%; Freq. No. 2,46 43;
35 Matches: 84; Conservative: 0; Mismatches: 0; Indels: 0; Caps:
36
37 1 FAYGQINRIVSASKEETPRFVVEELAEVFLVEVHQEVTECFPK 60
38 8 FALMCEHNFSSASAEHSTFYVCHDELDFELNVLAVVGVGVVGGHFN 67
39
40 61 LKRVCAHFLDMDPTFGHN 84
41
42 68 DDKVSAHPELMDPTFGHN 91
43
44 RESULT 13
45 US-08-480-478-50
46 Sequence 50; Application US/09480478
47 Patent No. 5864009
48 CHEMICAL INFORMATION:
49 APPLICANT: GEORGE P. VASARI; PATRICK ERIC
50 APPLICANT: HUGO STANSINS; JORIS HILDA
51 APPLICANT: LIVER NISSERS; MAG JOZEF
52 APPLICANT: LAURENTS; YVES RENE LAROCHE;
53 APPLICANT: LAURENT STEPHANE JESPER; and
54 APPLICANT: VANNICK GEORGES JOZEF
55 APPLICANT: GANSIEMANS
56 TITLE OF INVENTION: REMOVABLE EXTRACTED ANTI-
57 TITLE OF INVENTION: CHAMANT PROTEIN
58 NUMBER OF SEQUENCES: 86
59 ADDRESS/AGENT ADDRESS:
60 ADDRESSEE: Iyan & Iyan
61 STREET: 615 West Fifth Street
62 CITY: Suite 4700
63 CITY: Los Angeles
64 STATE: California
65 COUNTRY: U.S.A.
66 ZIP: 90071
67
68 COMPUTER READABLE FORM:
69 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
70 MEDIUM TYPE: storage
71 MEDIUM TYPE: COMPATIBLE

```

OPERATING SYSTEM: IBM P.C., DOS 5.0
 SOFTWARE: FAST SEQ VERSION 1.5
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: 08/26/1994
 FILING DATE: OCTOBER 1994
 CLASSIFICATION: 540
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/26/1994
 FILING DATE: 18 OCTOBER 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: BLOOM, SUZANNE L.
 REGISTRATION NUMBER: 40,158
 REFERENCE NUMBER: 200/226
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (214) 489-1600
 TELEFAX: (214) 955-0440
 TELEX: 67-4510
 INFORMATION FOR SEQ ID NO: 903
 SEQUENCE CHARACTERISTICS:
 LENGTH: 91 amino acids
 TYPE: amino acid
 COMPOSITION: 11000
 MISC: 400-476-50

Query Match 100.0% Score 4867 DB Z: Length 91
 Best Local Similarity 100.0% Ident. No. 2 to 13
 Matches 84: Conservative 02 Mismatches 02 Indels 02 Gaps 02

QY 1 FASTA: >US-09-498-556-59-128
 DB 8 KALP: >KALP-09-498-556-59-128
 DB 68 DOKVSAH: >DOKVSAH-09-498-556-59-128

QY 61 DOKVSAH: >DOKVSAH-09-498-556-59-128
 DB 68 DOKVSAH: >DOKVSAH-09-498-556-59-128

RESULT 14
 US-09-498-556-59-128
 Sequence 128: Application US/08486497
 Patent No. 586543
 GENERAL INFORMATION:
 APPLICANT: George P. Vlasuk, Patricia H. Stanssens,
 APPLICANT: Joris H. Meuniers, Marc J. Lauwereys,
 APPLICANT: Yves R. Laroche, Laurent S. Josephs,
 APPLICANT: Yandick G.L. Ganssmus, Matthew Molyto,
 APPLICANT: Peter W. Bergum
 TITLE OF INVENTION: RECOMBINANT EXTRACTED ANTICODONANT
 NUMBER OF SEQUENCES: 607
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: 644 West Fifth Street
 STREET: Suite 4700
 CITY: Los Angeles
 STATE: California
 COUNTRY: U.S.A.
 ZIP: 90071
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 MB
 MEDIUM TYPE: Storage
 COMPUTER: IBM compatible
 OPERATING SYSTEM: IBM PC, DOS 5.0
 SOFTWARE: Word Perfect 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: 08/26/1994
 FILING DATE: June 5, 1995
 CLASSIFICATION: 540
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/26/1994
 FILING DATE: October 18, 1994
 ATTORNEY/AGENT INFORMATION:

NAME: BLOOM, SUZANNE L.
 REGISTRATION NUMBER: 40,158
 REFERENCE NUMBER: 200/226
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (214) 489-1600
 TELEFAX: (214) 955-0440
 TELEX: 67-4510
 INFORMATION FOR SEQ ID NO: 128
 SEQUENCE CHARACTERISTICS:
 LENGTH: 91 amino acids
 TYPE: amino acid
 COMPOSITION: 11000
 MISC: 400-476-50

Query Match 100.0% Score 4867 DB Z: Length 91
 Best Local Similarity 100.0% Ident. No. 2 to 43
 Matches 84: Conservative 02 Mismatches 02 Indels 02 Gaps 02

QY 1 FASTA: >US-09-498-556-59-128
 DB 8 KALP: >KALP-09-498-556-59-128
 DB 68 DOKVSAH: >DOKVSAH-09-498-556-59-128

QY 61 DOKVSAH: >DOKVSAH-09-498-556-59-128
 DB 68 DOKVSAH: >DOKVSAH-09-498-556-59-128

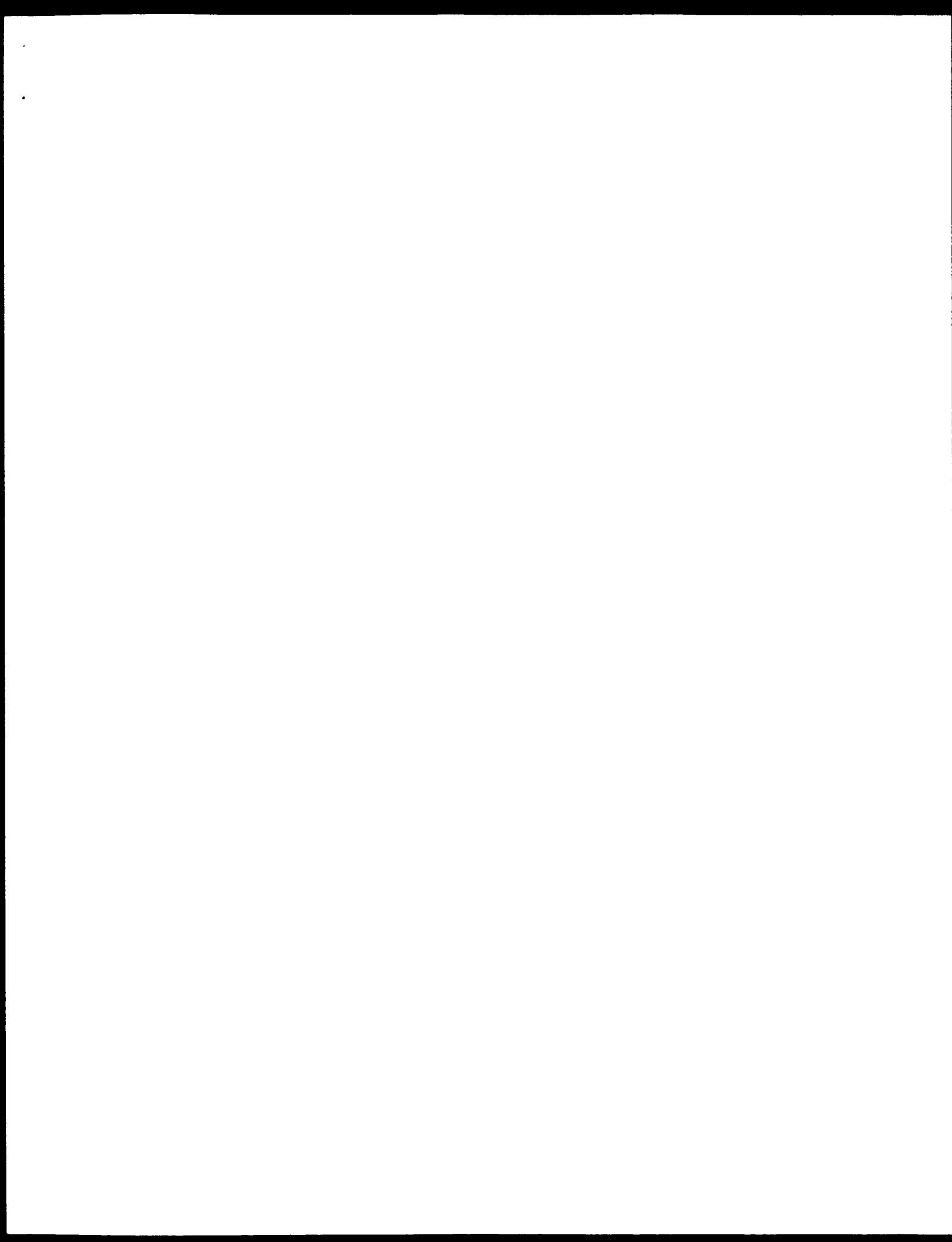
RESULT 15
 US-09-498-556-59-128
 Sequence 128: Application US/08486497
 Patent No. 586543
 GENERAL INFORMATION:
 APPLICANT: George P. Vlasuk, Patricia H. Stanssens,
 APPLICANT: Joris H. Meuniers, Marc J. Lauwereys,
 APPLICANT: Yves R. Laroche, Laurent S. Josephs,
 APPLICANT: Yandick G.L. Ganssmus, Matthew Molyto,
 APPLICANT: Peter W. Bergum
 TITLE OF INVENTION: RECOMBINANT EXTRACTED ANTICODONANT
 NUMBER OF SEQUENCES: 356
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: 644 West Fifth Street
 STREET: Suite 4700
 CITY: Los Angeles
 STATE: California
 COUNTRY: U.S.A.
 ZIP: 90071
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 MB
 MEDIUM TYPE: Storage
 COMPUTER: IBM compatible
 OPERATING SYSTEM: IBM PC, DOS 5.0
 SOFTWARE: Word Perfect 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: 08/26/1994
 FILING DATE: June 5, 1995
 CLASSIFICATION: 540
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/26/1994
 FILING DATE: October 18, 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: BLOOM, SUZANNE L.
 REGISTRATION NUMBER: 40,158
 REFERENCE NUMBER: 200/226
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (214) 489-1600
 TELEFAX: (214) 955-0440
 TELEX: 67-4510

1 INFORMATION FOR SEQ ID NO: 128:
2 SOURCE CHARACTERISTICS:
3 LENGTH: 91 amino acids
4 TYPE: amino acid
5 TOPOLOGY: 1 loop
6 MODEL TYPE: Peptide
7 ORIGINAL SOURCE:
8 ORGANISM: Anrylostoma caninum
9 US-09-498-556-128

Query Match: 100.0%; Score: 486; DB 2; Length: 91;
Best Local Similarity: 100.0%; Prod. No. 2.4e-43;
Matches: 84; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

QY 1 KATWQVFNPVQVNSFTQKPKVTVVPEPELENVPELVVWVHQAQVPEVPEVNR 60
1b 8 KATWQVFNPVQVNSFTQKPKVTVVPEPELENVPELVVWVHQAQVPEVPEVNR 67
QY 61 DDKVSAEDPELNMDETYPIRN 84
1b 68 DDKVSAEDPELNMDETYPIRN 91

Search completed: September 30, 2001, 10:29:37
Job time: 47 sec




```

14 01 MAY 2001 (TREMBL:16, last sequence update)
15 01 MAR 2001 (TREMBL:16, last annotation update)
16 Yeast 48 PROTEIN.
17 Yeast 48.
18 Caenorhabditis elegans.
19 Eukaryota: Metazoa: Nemertea: Chromadorea: Rhadaliidae: Rhadaliidae?
20 Rhadaliidae: Polychaeta: Caprellidae.
21 NCBI Taxid 62593.
22 111.
23 SEQUENCE FROM N.A.
24 BARNETT A.A.
25 1.2.
26 1.2.
27 SEQUENCE FROM N.A.
28 SEQUENCE FROM N.A.
29 SEQUENCE FROM N.A.
30 SEQUENCE FROM N.A.
31 SEQUENCE FROM N.A.
32 SEQUENCE FROM N.A.
33 SEQUENCE FROM N.A.
34 SEQUENCE FROM N.A.
35 SEQUENCE FROM N.A.
36 SEQUENCE FROM N.A.
37 SEQUENCE FROM N.A.
38 SEQUENCE FROM N.A.
39 SEQUENCE FROM N.A.
40 SEQUENCE FROM N.A.
41 SEQUENCE FROM N.A.
42 SEQUENCE FROM N.A.
43 SEQUENCE FROM N.A.
44 SEQUENCE FROM N.A.
45 SEQUENCE FROM N.A.
46 SEQUENCE FROM N.A.
47 SEQUENCE FROM N.A.
48 SEQUENCE FROM N.A.
49 SEQUENCE FROM N.A.
50 SEQUENCE FROM N.A.
51 SEQUENCE FROM N.A.
52 SEQUENCE FROM N.A.
53 SEQUENCE FROM N.A.
54 SEQUENCE FROM N.A.
55 SEQUENCE FROM N.A.
56 SEQUENCE FROM N.A.
57 SEQUENCE FROM N.A.
58 SEQUENCE FROM N.A.
59 SEQUENCE FROM N.A.
60 SEQUENCE FROM N.A.
61 SEQUENCE FROM N.A.
62 SEQUENCE FROM N.A.
63 SEQUENCE FROM N.A.
64 SEQUENCE FROM N.A.
65 SEQUENCE FROM N.A.
66 SEQUENCE FROM N.A.
67 SEQUENCE FROM N.A.
68 SEQUENCE FROM N.A.
69 SEQUENCE FROM N.A.
70 SEQUENCE FROM N.A.
71 SEQUENCE FROM N.A.
72 SEQUENCE FROM N.A.
73 SEQUENCE FROM N.A.
74 SEQUENCE FROM N.A.
75 SEQUENCE FROM N.A.
76 SEQUENCE FROM N.A.
77 SEQUENCE FROM N.A.
78 SEQUENCE FROM N.A.
79 SEQUENCE FROM N.A.
80 SEQUENCE FROM N.A.
81 SEQUENCE FROM N.A.
82 SEQUENCE FROM N.A.
83 SEQUENCE FROM N.A.
84 SEQUENCE FROM N.A.
85 SEQUENCE FROM N.A.
86 SEQUENCE FROM N.A.
87 SEQUENCE FROM N.A.
88 SEQUENCE FROM N.A.
89 SEQUENCE FROM N.A.
90 SEQUENCE FROM N.A.
91 SEQUENCE FROM N.A.
92 SEQUENCE FROM N.A.
93 SEQUENCE FROM N.A.
94 SEQUENCE FROM N.A.
95 SEQUENCE FROM N.A.
96 SEQUENCE FROM N.A.
97 SEQUENCE FROM N.A.
98 SEQUENCE FROM N.A.
99 SEQUENCE FROM N.A.
100 SEQUENCE FROM N.A.

```

```

14 01 MAY 2001 (TREMBL:16, last sequence update)
15 01 MAR 2001 (TREMBL:16, last annotation update)
16 Yeast 48 PROTEIN.
17 Yeast 48.
18 Caenorhabditis elegans.
19 Eukaryota: Metazoa: Nemertea: Chromadorea: Rhadaliidae: Rhadaliidae?
20 Rhadaliidae: Polychaeta: Caprellidae.
21 NCBI Taxid 62593.
22 111.
23 SEQUENCE FROM N.A.
24 BARNETT A.A.
25 1.2.
26 1.2.
27 SEQUENCE FROM N.A.
28 SEQUENCE FROM N.A.
29 SEQUENCE FROM N.A.
30 SEQUENCE FROM N.A.
31 SEQUENCE FROM N.A.
32 SEQUENCE FROM N.A.
33 SEQUENCE FROM N.A.
34 SEQUENCE FROM N.A.
35 SEQUENCE FROM N.A.
36 SEQUENCE FROM N.A.
37 SEQUENCE FROM N.A.
38 SEQUENCE FROM N.A.
39 SEQUENCE FROM N.A.
40 SEQUENCE FROM N.A.
41 SEQUENCE FROM N.A.
42 SEQUENCE FROM N.A.
43 SEQUENCE FROM N.A.
44 SEQUENCE FROM N.A.
45 SEQUENCE FROM N.A.
46 SEQUENCE FROM N.A.
47 SEQUENCE FROM N.A.
48 SEQUENCE FROM N.A.
49 SEQUENCE FROM N.A.
50 SEQUENCE FROM N.A.
51 SEQUENCE FROM N.A.
52 SEQUENCE FROM N.A.
53 SEQUENCE FROM N.A.
54 SEQUENCE FROM N.A.
55 SEQUENCE FROM N.A.
56 SEQUENCE FROM N.A.
57 SEQUENCE FROM N.A.
58 SEQUENCE FROM N.A.
59 SEQUENCE FROM N.A.
60 SEQUENCE FROM N.A.
61 SEQUENCE FROM N.A.
62 SEQUENCE FROM N.A.
63 SEQUENCE FROM N.A.
64 SEQUENCE FROM N.A.
65 SEQUENCE FROM N.A.
66 SEQUENCE FROM N.A.
67 SEQUENCE FROM N.A.
68 SEQUENCE FROM N.A.
69 SEQUENCE FROM N.A.
70 SEQUENCE FROM N.A.
71 SEQUENCE FROM N.A.
72 SEQUENCE FROM N.A.
73 SEQUENCE FROM N.A.
74 SEQUENCE FROM N.A.
75 SEQUENCE FROM N.A.
76 SEQUENCE FROM N.A.
77 SEQUENCE FROM N.A.
78 SEQUENCE FROM N.A.
79 SEQUENCE FROM N.A.
80 SEQUENCE FROM N.A.
81 SEQUENCE FROM N.A.
82 SEQUENCE FROM N.A.
83 SEQUENCE FROM N.A.
84 SEQUENCE FROM N.A.
85 SEQUENCE FROM N.A.
86 SEQUENCE FROM N.A.
87 SEQUENCE FROM N.A.
88 SEQUENCE FROM N.A.
89 SEQUENCE FROM N.A.
90 SEQUENCE FROM N.A.
91 SEQUENCE FROM N.A.
92 SEQUENCE FROM N.A.
93 SEQUENCE FROM N.A.
94 SEQUENCE FROM N.A.
95 SEQUENCE FROM N.A.
96 SEQUENCE FROM N.A.
97 SEQUENCE FROM N.A.
98 SEQUENCE FROM N.A.
99 SEQUENCE FROM N.A.
100 SEQUENCE FROM N.A.

```
